

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



56

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : C12Q 1/34, C12N 9/24, 9/96, C07K 14/47		A1	(11) International Publication Number: WO 95 04158
			(43) International Publication Date: 9 February 1995 (09.02.95)
(21) International Application Number: PCT/US94/08207		(74) Agent: JAMESON, William, G.; The Upjohn Company, Corporate Intellectual Property Law, 301 Henrietta Street, Kalamazoo, MI 49001 (US).	
(22) International Filing Date: 26 July 1994 (26.07.94)			
(30) Priority Data: 08/099,866 29 July 1993 (29.07.93) US 08/136,117 13 October 1993 (13.10.93) US		(81) Designated States: AM, AT, AU, BB, BG, BR, BY, CA, CH, CN, CZ, DE, DK, ES, FI, GB, GE, HU, JP, KE, KG, KP, KR, KZ, LK, LT, LU, LV, MD, MG, MN, MW, NL, NO, NZ, PL, PT, RO, RU, SD, SE, SI, SK, TJ, TT, UA, US, UZ, VN, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG), ARIPO patent (KE, MW, SD).	
(60) Parent Application or Grant (63) Related by Continuation US 08/136,117 (CIP) Filed on 13 October 1993 (13.10.93)		Published With international search report. Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.	
(71) Applicant (for all designated States except US): THE UPJOHN COMPANY [US/US]; 301 Henrietta Street, Kalamazoo, MI 49001 (US).			
(72) Inventors; and (75) Inventors/Applicants (for US only): HOOGWERF, Arlene, J. [US/US]; 8435 Bluebird, Richland, MI 49083 (US). LEDBETTER, Steven, R. [US/US]; 550 West Bridge Street, Plainwell, MI 49080 (US).			
(54) Title: USE OF HEPARANASE TO IDENTIFY AND ISOLATE ANTI-HEPARANASE COMPOUND			
(57) Abstract  Purified heparanase having activity of greater than 20 units/ $\mu$ g protein, preferably greater than 50 units heparanase activity per $\mu$ g protein, is described. The use of heparanase for screening for anti-heparanase compounds is also described. In addition, the use of the high potency heparanase to accelerate wound healing or its use as an immobilized heparanase filter connected to extracorporeal devices to degrade heparin and neutralize its anticoagulant properties during surgery is disclosed.			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	GB	United Kingdom	MR	Mauritania
AU	Australia	GE	Georgia	MW	Malawi
BB	Barbados	GN	Guinea	NE	Niger
BE	Belgium	GR	Greece	NL	Netherlands
BF	Burkina Faso	HU	Hungary	NO	Norway
BG	Bulgaria	IE	Ireland	NZ	New Zealand
BJ	Benin	IT	Italy	PL	Poland
BR	Brazil	JP	Japan	PT	Portugal
BY	Belarus	KE	Kenya	RO	Romania
CA	Canada	KG	Kyrgyzstan	RU	Russian Federation
CF	Central African Republic	KP	Democratic People's Republic of Korea	SD	Sudan
CG	Congo	KR	Republic of Korea	SE	Sweden
CH	Switzerland	KZ	Kazakhstan	SI	Slovenia
CI	Côte d'Ivoire	LI	Liechtenstein	SK	Slovakia
CM	Cameroon	LK	Sri Lanka	SN	Senegal
CN	China	LU	Luxembourg	TD	Chad
CS	Czechoslovakia	LV	Latvia	TG	Togo
CZ	Czech Republic	MC	Monaco	TJ	Tajikistan
DE	Germany	MD	Republic of Moldova	TT	Trinidad and Tobago
DK	Denmark	MG	Madagascar	UA	Ukraine
ES	Spain	ML	Mali	US	United States of America
FI	Finland	MN	Mongolia	UZ	Uzbekistan
FR	France			VN	Viet Nam
GA	Gabon				

-1-

USE OF HEPARANASE TO IDENTIFY AND ISOLATE ANTI-HEPARANASE  
COMPOUND

FIELD OF INVENTION

The present invention discloses the use of mammalian heparanase, preferably  
5 recombinant heparanase, for screening for anti-heparanase compounds. More particularly, the  
present invention provides a method of selecting IHA (Inhibitors of Heparanase Activity). In  
addition, the present invention provides a purified heparanase, particularly suitable for use to  
identify and isolate anti-heparanase compounds as well as for other known uses of heparanases,  
such as its use to accelerate wound healing or its use as an immobilized heparanase filter  
10 connected to extracorporeal devices to degrade heparin and neutralize its anticoagulant properties  
during surgery.

BACKGROUND OF THE INVENTION

Elevated heparanase activity has been documented in mobile, invasive cells. Examples  
include: invasive melanoma, lymphoma, mastocytoma, mammary adenocarcinoma, leukemia,  
15 and rheumatoid fibroblasts. This activity has also been documented in non-pathologic  
situations involving the migration of lymphocytes, neutrophils, macrophages, eosinophils and  
platelets. An inhibitor of heparanase would therefore broadly influence the invasive potential of  
these diverse cells.

Inhibition of heparan sulfate degradation would also inhibit the release of bound growth  
20 factors and other biologic response modifiers that would, if released, fuel the growth of adjacent  
tissues and provide a supportive environment for cell growth (Rapraeger, et al., *Science* 252:  
1705-1708, 1991). Inhibitors of heparanase activity would be of value in the treatment of  
arthritis, vascular restenosis, tumor growth and progression, and fibro-proliferative disorders.

Until now, the obstacles to designing a screening assay to find inhibitors of mammalian  
25 heparanase have been the unavailability of a mammalian heparanase that is purified to apparent  
homogeneity and the lack of information about the amino acid sequence or the 3-dimensional  
structure of the enzyme. Without the amino acid sequence, it has not been possible to produce  
recombinant mammalian heparanase to be used in large volume screening efforts. Knowledge  
of the tertiary and quaternary structures would facilitate rational design of IHA. This report  
30 overcomes obstacles relating to the sequence of the heparanase, and also provides a model for  
higher-order structure.

Heparanase refers to a mammalian enzyme which can degrade heparin proteoglycans  
(HPG) and/or heparan sulfate proteoglycans (HSPG).

Heparanase activity in mammalian cells is well known. It is found in various melanoma  
35 cells (Nakajima, et al., *Cancer Letters* 31: 277-283, 1986), mammary adenocarcinoma cells  
(Parish, et al., *Int. J. Cancer*, 40: 511-518, 1987), leukemic cells (Yahalom, et al., *Leukemia*

*Research* 12: 711-717, 1988), mast cells (Ogren and Lindahl, *J. Biol. Chem.* 250: 2690-2697, 1975), macrophages (Savion, *et al.*, *J. Cell. Physiol.*, 130: 85-92, 1987), mononuclear cells (Sewell, *et al.*, *Biochem. J.* 264: 777-783, 1989), neutrophils (Matzner, *et al.*, *J. Leukocyte Biology* 51: 519-524, 1992), T-cells (Vettel, *et al.*, *Eur. J. Immunol.* 21: 2247-2251, 1991),  
5 platelets (Haimovitz-Friedman, *et al.*, *Blood* 78: 789-796, 1991), endothelial cells (Godder, *et al.*, *J. Cell Physiol.* 148: 274-280, 1991), and placenta (Klein and von Figura, *BBRC* 73: 569, 1976).

WO 91/02977, incorporated herein by reference, describes a substantially, but partially, purified heparanase produced by cation exchange resin chromatography and the affinity  
10 absorbent purification of heparanase-containing cell extract. WO 91/02977 also describes a method promoting wound healing utilizing compositions comprising a "purified" form of heparanase.

Others have proposed the use of purified bacterial heparanase, immobilized onto filters and connected to extracorporeal devices, to degrade heparin and neutralize its anticoagulant  
15 properties post surgery (Freed, *et al.*, *Ann. Biomed. Eng.* 21: 67-76, 1993).

U.S. Patent 4,882,318 describes heparanase-inhibiting compositions for preventing tumor metastasis.

Haimovitz-Friedman *et al.* (*Blood* 78: 789-796, 1991) describe an assay for heparanase activity that involves the culturing of endothelial cells in radiolabeled  $^{35}\text{SO}_4$  to produce  
20 radiolabeled heparan sulfate proteoglycans, the removal of the cells which leaves the deposited extracellular matrix that contains the  $^{35}\text{S}$ -HSPG, the addition of potential sources of heparanase activity, and the detection of possible activity by passing the supernatant from the radiolabeled extracellular matrix over a gel filtration column and monitoring for changes of the size of the radiolabeled material that would indicate that HSPG degradation had taken place. This assay  
25 does not have the capability for large-scale screening of inhibitors.

Nakajima *et al.* (*Anal. Biochem.* 196: 162-171, 1986) describe a solid-phase substrate for the assay of melanoma heparanase activity. Heparan sulfate from bovine lung is chemically radiolabeled by reacting it with  $[^{14}\text{C}]$ -acetic anhydride. Free amino groups of the  $[^{14}\text{C}]$ -heparan sulfate were acetylated and the reducing termini were aminated. The  $[^{14}\text{C}]$ -heparan sulfate was  
30 chemically coupled to an agarose support via the introduced amine groups on the reducing termini. This substrate is limited in that it is an extensively chemically modified form of naturally occurring heparan sulfate.

Khan and Newman (*Anal. Biochem.* 196: 373-376, 1991) describe an indirect assay for heparanase activity. In this assay, heparin is quantitated by its ability to interfere with the color  
35 development between a protein and the dye Coomassie brilliant blue. Heparanase activity is

detected by the loss of this interference. This assay is limited in use for screening because it is so indirect that other non-heparin compounds could also interfere with the protein-dye reaction.

The CXC chemokine family (also called the intercrine  $\alpha$  family) is one branch of the supergene "intercrine" cytokine family (Oppenheim, *Ann. Rev. Biochem.* 9: 617-648, 1991).

- 5 It's members include platelet factor 4, platelet basic protein and derivatives,  $\gamma$ IP-10, gro( $\alpha, \beta, \gamma$ ), NAP-1/interleukin-8, mig, and ENA-78 (for review, see Miller and Krangel, *Critical Reviews in Immunology* 12: 17-46, 1992). The other branch, the CC chemokines or intercrine- $\beta$  family, includes MIP1 $\alpha$ , MIP1 $\beta$ , JE/MCP-1, RANTES, and MCAF. All members of both branches of this chemokine family characteristically are basic heparin-binding polypeptides, display
- 10 molecular weights between 8 and 11 kD, share 20 - 50% homology, and function broadly in pathologic situations characterized by inflammation and tissue remodeling.

- The proteolytically processed forms of platelet basic protein include CTAP-III,  $\beta$ -thromboglobulin, and NAP-2.  $\beta$ -thromboglobulin (Moore, *et al.*, *Biochim. Biophys. Acta.* 379: 360-369, 1975) and CTAP-III (Castor, *et al.*, *Arthritis Rheum.* 20: 859-868, 1977), were
- 15 originally isolated from activated supernatants or lysates from outdated platelets. Using the techniques of subcellular fractionation and radioimmunoassay,  $\beta$ -thromboglobulin was identified as an  $\alpha$ -granule protein that could be released upon activation (Kaplan, *et al.*, *Blood* 53: 604-618, 1979). Platelet basic protein itself was later isolated from fresh platelets, megakaryocytes, and HEL cells, an immortal human erythroleukemia cell line (Holt, *et al.*, *Biochemistry* 25:
- 20 1988-1996, 1986; Holt, *et al.*, *Exp. Hematol.* 16: 302-306, 1988). Walz and Baggiolini isolated the processed form of NAP-2 from platelet-containing cultures of stimulated mononuclear cells (Walz, *et al.*, *J. Exp. Med.* 170: 1745-1750, 1989).

- Material labeled as  $\beta$ -thromboglobulin is commercially available from Calbiochem, San Diego, CA (Cat. # 605165), Celsus Laboratories, Cincinnati, OH (Cat. # 41705), and
- 25 Haematologic Technologies, Essex Jct., VT (Cat. # HBTG-02100). The inventors have determined, by using the "Purification Assay," that the commercial preparation have heparanase activity at a level of 0.075 units/ $\mu$ g. This activity is below the level of 1 unit/ $\mu$ g needed for the screening of anti-heparanase compounds in accordance with the assay of the subject invention.

- U.S. Patent 4,897,348 describes recombinant materials and methods for producing
- 30 human connective tissue-activating peptide-III (CTAP-III) and analogs thereof.

- Transglutaminases catalyze the posttranslational modification of proteins by transamidation of available glutamine residues. This action results primarily in the formation of epsilon-(gamma-glutamyl)lysine cross-links (Greenberg, *et al.*, *FASEB J.* 5: 3071-3076, 1991). This posttranslational modification has been reported to dramatically alter the action of some
- 35 small proteins. For example, a transglutaminase produces a glutamine-lysine cross-link in the 13

kD phospholipase A<sub>2</sub> and increases its specific enzymatic activity (Cordella-Miele, *et al.*, *J. Biol. Chem.* **265**: 17180-17188, 1990). A transglutaminase cross-links another small molecule, interleukin-2, and converts its activity to one that is cytotoxic to mature oligodendrocytes (Eitan and Schwartz, *Science* **261**: 106-108, 1993). The glutamine-lysine cross-link in a protein would result in the loss of overall positive charge for that protein. The transglutaminases are optimally active and generally used under reducing conditions such as dithiothreitol. The concept that glutamine-lysine cross-linking alters the activity of these small proteins may be applicable to other small molecules as well.

#### SUMMARY OF THE INVENTION

The present invention discloses a method of screening for compounds having anti-heparanase activity (AHA compounds), i.e. inhibitors of heparanase activity (IHA), comprising the steps of: contacting a potential AHA compound with radiolabeled heparin/heparan sulfate and heparanase for a time and under such conditions sufficient to allow for inhibition of heparanase activity; detecting inhibition of heparanase activity; and selecting compounds that inhibit heparanase activity. The present invention also discloses the amino acid sequence identity of the heparanase that has been purified to homogeneity by chromatography under reducing conditions. Identification of the amino acid sequence of the protein which contains heparanase activity is crucial for the production of recombinant mammalian heparanase.

#### DETAILED DESCRIPTION OF THE INVENTION

The present invention provides a purified heparanase, and a method for producing it. The heparanase so produced has an activity of greater than 20 units/  $\mu$ g protein, preferably greater than 50 units heparanase activity per  $\mu$ g protein (1 unit = 1% cpm < 30 K/hr using the "Purification Assay" (Example 2, Part D).

In addition, the present invention provides recombinant heparanase and a means for producing it. The term "purified heparanase" as used in the specification and claims includes the recombinant heparanase as described in the subject application. The recombinant heparanase of the subject invention can be used for the same purposes and in the same manner as the purified heparanase.

The purified heparanase of the present invention has an isoelectric point of less than 5.5 (preferably about 4.8 - 5.1) and preferably is activated by treatment with transglutaminase using reducing conditions.

The recombinant heparanase of the present invention has an isoelectric point of less than 5.5 (preferably about 4.8 - 5.1), and is isolated under reducing conditions and is activated by treatment with transglutaminase.

Suitable transglutaminases that may be used for this purpose include Activated Factor XIIIa, guinea pig liver transglutaminase, epidermal transglutaminase, keratinocyte

-5-

transglutaminase, and tissue transglutaminase.

The heparanase of the present invention has the amino acid sequence (SEQ. ID. NO: 1) of:

```

5   Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala
      5           10           15

    Glu Leu Arg Cys Met Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys
      20           25           30

10  Asn Ile Gln Ser Leu Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln
      35           40           45

    Val Glu Val Ile Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp
      50           55           60

15  Pro Asp Ala Pro Arg Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly
      65           70           75           80

    Asp Glu Ser Ala Asp
      85
  
```

encoded by the cDNA sequence (SEQ ID NO: 2) of:

```

1  AACTTGGCGA AAGGCAAAGA GGAAAGTCTA GACAGTGACT TGTATGCTGA
51 ACTCCGCTGC ATGTGTATAA AGACAACCTC TGGAATTCAT CCCAAAAACA
101 TCCAAAGTTT GGAAGTGATC GGGAAAGGAA CCCATTGCAA CCAAGTCGAA
25 151 GTGATAGCCA CACTGAAGGA TGGGAGGAAA ATCTGCCTGG ACCCAGATGC
201 TCCCAGAATC AAGAAAATTG TACAGAAAAA ATTGGCAGGT GATGAATCTG
251 CTGAT
  
```

which corresponds to the cDNA sequence and derived amino acid sequence of CTAP-III. See Wenger et al., *Blood*, 73: 1498-1503, 1989.

30 In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 3) of:

```

    Ser Ser Thr Lys Gly Gln Thr Lys Arg Asn Leu Ala Lys Gly Lys Glu
      5           10           15

35  Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu Leu Arg Cys Met Cys Ile
      20           25           30

    Lys Thr Thr Ser Gly Ile His Pro Lys Asn Ile Gln Ser Leu Glu Val
      35           40           45

40  Ile Gly Lys Gly Thr His Cys Asn Gln Val Glu Val Ile Ala Thr Leu
      50           55           60

    Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro Arg Ile Lys
45 65           70           75           80
  
```

-6-

Lys Ile Val Gln Lys Lys Leu Ala Gly Asp Glu Ser Ala Asp  
85 90

encoded by the cDNA sequence (SEQ ID NO: 4) of:

1 TCCTCCACCA AAGGACAAAC TAAGAGAAAC TTGGCGAAAG GCAAAGAGGA  
5 51 AAGTCTAGAC AGTGACTTGT ATGCTGAACT CCGCTGCATG TGTATAAAGA  
101 CAACCTCTGG AATTCATCCC AAAAACATCC AAAGTTTGGA AGTGATCGGG  
151 AAAGGAACCC ATTGCAACCA AGTCGAAGTG ATAGCCACAC TGAAGGATGG  
201 GAGGAAAATC TGCCTGGACC CAGATGCTCC CAGAATCAAG AAAATTGTAC  
251 AGAAAAAATT GGCAGGTGAT GAATCTGCTG AT

10 which corresponds to the cDNA sequence and derived amino acid sequence of platelet basic protein. See Wenger et al., *Blood*, 73: 1498-1503, 1989 as well as Walz and Baggiolini, *BBRC* 159: 969-981, 1989; Castor, et al., *BBRC* 163: 1071-1078, 1989.

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 5) of:

15 Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu Leu Arg Cys  
1 5 10 15

Met Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys Asn Ile Gln Ser  
20 25 30

20 Leu Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln Val Glu Val Ile  
35 40 45

25 Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro  
50 55 60

Arg Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly Asp Glu Ser Ala  
65 70 75 80

30 Asp

encoded by the cDNA sequence (SEQ ID NO: 6) of:

1 GGCAAAGAGG AAAGTCTAGA CAGTGACTTG TATGCTGAAC TCCGCTGCAT  
51 GTGTATAAAG ACAACCTCTG GAATTCATCC CAAAAACATC CAAAGTTTGG  
101 AAGTGATCGG GAAAGGAACC CATTGCAACC AAGTCGAAGT GATAGCCACA  
35 151 CTGAAGGATG GGAGGAAAAT CTGCCTGGAC CCAGATGCTC CCAGAATCAA  
201 GAAAATTGTA CAGAAAAAAT TGGCAGGTGA TGAATCTGCT GAT

which corresponds to the cDNA sequence and derived amino acid sequence  $\beta$ -thromboglobulin ( $\beta$ TG). See Wenger et al., *Blood*, 73: 1498-1503, 1989.

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 7) of:



-7-

Glu Leu Arg Cys Met Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys  
 1                      5                      10                      15  
 Asn Ile Gln Ser Leu Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln  
 5                      20                      25                      30  
 Val Glu Val Ile Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp  
                     35                      40                      45  
 10 Pro Asp Ala Pro Arg Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly  
                     50                      55                      60  
 Asp Glu Ser Ala Asp  
 65  
 15

encoded by the cDNA sequence (SEQ ID NO: 8) of:

1 GAACTCCGCT GCATGTGTAT AAAGACAACC TCTGGAATTC ATCCCAAAAA  
 51 CATCCAAAGT TTGGAAGTGA TCGGGAAAGG AACCCATTGC AACCAAGTCG  
 101 AAGTGATAGC CACACTGAAG GATGGGAGGA AAATCTGCCT GGACCCAGAT  
 20 151 GCTCCCAGAA TCAAGAAAAT TGTACAGAAA AAATTGGCAG GTGATGAATC  
 201 TGCTGAT

which corresponds to the cDNA sequence and derived amino acid sequence of neutrophil activating peptide-2.

The foregoing amino acid sequences correspond to the products of a single gene called  
 25 platelet basic protein (Walz and Baggiolini, *BBRC* 159: 969-981, 1989; Castor, *et al.*, *BBRC*  
 163: 1071-1078, 1989). The complete gene sequence of platelet basic protein is well known.  
 See, for example, Wenger *et al.*, *Blood*, 73: 1498-1503, 1989 and *Proc. Natl. Acad. Sci. USA*,  
 90, 3660-3664, 1993.

The present invention also provides heparanase having the amino acid sequences of  
 30 other members of the CXC chemokine family [including Platelet factor 4 (SEQ. ID NO. 12),  
 $\gamma$ IP-10 (SEQ. ID NO. 14), *gro*/MGSA (SEQ. ID NO. 16), *gro*- $\beta$ /MIP-2 $\alpha$  (SEQ. ID NO. 18),  
*gro*- $\gamma$ /MIP-2 $\beta$  (SEQ. ID NO. 20), Interleukin-8/NAP-1 (SEQ. ID NO. 22) and ENA-78 (SEQ. ID  
 NO. 24)] as well as members of the CC chemokine family [including MIP-1 $\alpha$  (SEQ. ID NO.  
 26), MIP-1 $\beta$  (SEQ. ID NO. 28), I-309 (SEQ. ID NO. 23), MCP-1 (SEQ. ID NO. 32), MCP-3  
 35 (SEQ. ID NO. 34), RANTES (SEQ. ID NO. 36), fic (SEQ. ID NO. 38) and MCP-2 (SEQ. ID  
 NO. 40)]; purified to apparent homogeneity, prepared in the presence of reducing conditions,  
 and activated by treatment with transglutaminase. Suitable transglutaminases that may be used  
 for this purpose include Activated Factor XIIIa, guinea pig liver transglutaminase, epidermal  
 transglutaminase, keratinocyte transglutaminase, and tissue transglutaminase.

40 In another aspect, the present invention provides a heparanase having the amino acid

sequence (SEQ ID NO: 12) of:

Met Ser Ser Ala Ala Gly Phe Cys Ala Ser Arg Pro Gly Leu Leu Phe Leu Gly Leu Leu  
 Leu Leu Pro Leu Val Val Ala Phe Ala Ser Ala Glu Ala Glu Glu Asp Gly Asp Leu Gln  
 Cys Leu Cys Val Lys Thr Thr Ser Gln Val Arg Pro Arg His Ile Thr Ser Leu Glu Val  
 5 Ile Lys Ala Gly Pro His Cys Pro Thr Ala Gln Leu Ile Ala Thr Leu Lys Asn Gly Arg  
 Lys Ile Cys Leu Asp Leu Gln Ala Pro Leu Tyr Lys Lys Ile Ile Lys Lys Leu Leu Glu  
 Ser

encoded by the cDNA sequence (SEQ ID NO: 13) of:

1 CCGCAGCATG AGCTCCGCAG CCGGGTTCTG CGCCTCACGC CCCGGGCTGC  
 10 51 TGTTCTTGGG GTTGCTGCTC CTGCCACTTG TGGTCGCCTT CGCCAGCGCT  
 101 GAAGCTGAAG AAGATGGGGA CCTGCAGTGC CTGTGTGTGA AGACCACCTC  
 151 CCAGGTCCGT CCCAGGCACA TCACCAGCCT GGAGGTGATC AAGGCCGGAC  
 201 CCCACTGCCC CACTGCCCAA CTGATAGCCA CGCTGAAGAA TGGAAGGAAA  
 251 ATTTGCTTGG ACCTGCAAGC CCCGCTGTAC AAGAAAATAA TTAAGAAACT  
 15 301 TTTGGAGAGT TAGCTACTAG CTGCCTACGT GTGTGCATTT GCTATATAGC  
 351 ATACTTCTTT TTTCCAGTTT CAATCTAACT GTGAAAGAAA CTTCTGATAT  
 401 TTGTGTTATC CTTATGATTT TAAATAAACA AAATAAATC

which corresponds to the cDNA sequence and derived amino acid sequence of platelet factor 4.  
 See Poncz et al., *Blood* 69, 219-223 (1987).

20 In another aspect, the present invention provides a heparanase having the amino acid  
 sequence (SEQ ID NO: 14) of:

Met Asn Gln Thr Ala Ile Leu Ile Cys Cys Leu Ile Phe Leu Thr Leu Ser Gly Ile Gln Gly  
 Val Pro Leu Ser Arg Thr Val Arg Cys Thr Cys Ile Ser Ile Ser Asn Gln Pro Val Asn Pro  
 Val Asn Pro Arg Ser Leu Glu Lys Leu Glu Ile Ile Pro Ala Ser Gln Phe Cys Pro Arg  
 25 Val Glu Ile Ile Ala Thr Met Lys Lys Lys Gly Glu Lys Arg Cys Leu Asn Pro Glu Ser  
 Lys Ala Ile Lys Asn Leu Leu Lys Ala Val Ser Lys Glu Met Ser Lys Arg Ser Pro

encoded by the cDNA sequence (SEQ ID NO: 15) of:

1 GAGACATTCC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC  
 51 CTCCAGTCTC AGCACCATGA ATCAAAGTGC GATTCTGATT TGCTGCCTTA  
 30 101 TCTTTCTGAC TCTAAGTGGC ATTCAAGGAG TACCTCTCTC TAGAACCGTA  
 151 CGCTGTACCT GCATCAGCAT TAGTAATCAA CCTGTTAATC CAAGGTCTTT  
 201 AGAAAAACTT GAAATTATTC CTGCAAGCCA ATTTGTCCA CGTGTGAGA  
 251 TCATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA  
 301 TCGAAGGCCA TCAAGAATTT ACTGAAAGCA GTTAGCAAGG AAATGTCTAA  
 35 351 AAGATCTCCT TAAAACCAGA GGGGAGCAAA ATCGATGCAG TGCTTCCAAG  
 401 GATGGACCAC ACAGAGGCTG CCTCTCCCAT CACTCCCTA CATGGAGTAT

451 ATGTCAAGCC ATAATTGTTC TTAGTTTGCA GTTACACTAA AAGGTGACCA  
 501 ATGATGGTCA CCAAATCAGC TGCTACTACT CCTGTAGGAA GGTAAATGTT  
 551 CATCATCCTA AGCTATTAGC TAATAACTCT ACCCTGGCAC TATAATGTAA  
 601 GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTTCTGACC CTGCTTCAAA

- 5 which corresponds to the cDNA sequence and derived amino acid sequence  $\gamma$ IP-10. See Luster et al., Nature 315, 672-676 (1985).

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 16) of:

Met Ala Arg Ala Ala Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu Arg Val Ala Leu  
 10 Leu Leu Leu Leu Val Ala Ala Gly Arg Arg Ala Ala Gly Ala Ser Val Ala Thr Glu  
 Leu Arg Cys Gln Cys Leu Gln Thr Leu Gln Gly Ile His Pro Lys Asn Ile Gln Ser Val  
 Asn Val Lys Ser Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn  
 Gly Arg Lys Ala Cys Leu Asn Pro Ala Ser Pro Ile Val Lys Lys Ile Ile Glu Lys Met  
 Leu Asn Ser Asp Lys Ser Asn

- 15 encoded by the cDNA sequence (SEQ ID NO: 17) of:

1 CTCGCCAGCT CTTCCGCTCC TCTCACAGCC GCCAGACCCG CCTGCTGAGC  
 51 CCCATGGCCC GCGCTGCTCT CTCCGCCGCC CCCAGCAATC CCCGGCTCCT  
 101 GCGAGTGGCA CTGCTGCTCC TGCTOCTGGT AGCCGCTGGC CGGCGCGCAG  
 151 CAGGAGCGTC CGTGGCCACT GAACTGCGCT GCCAGTGCTT GCAGACCCTG  
 20 201 CAGGGAATTC ACCCCAAGAA CATCCAAAGT GTGAACGTGA AGTCCCCCGG  
 251 ACCCCACTGC GCCCAAACCG AAGTCATAGC CAACTCAAG AATGGGCGGA  
 301 AAGCTTGCCT CAATCCTGCA TCCCCATAG TTAAGAAAAT CATCGAAAAG  
 351 ATGCTGAACA GTGACAAATC CAACTGACCA GAAGGGAGGA GGAAGCTCAC  
 401 TGGTGGCTGT TCCTGAAGGA GGCCCTGCCC TTATAGGAAC AGAAGAGGAA  
 25 451 AGAGAGACAC AGCTGCAGAG GCCACCTGGA TTGTGCCTAA TGTGTTTGAG  
 501 CATCGCTTAG GAGAAGTCTT CTATTTATTT ATTTATTCAT TAGTTTTGAA  
 551 GATTCTATGT TAATATTTTA GGTGTAAAAT AATTAAGGGT ATGATTAAC  
 601 CTACCTGCAC ACTGTCCTAT TATATTCATT CTTTTTGAAA TGTCAACCCC  
 651 AAGTTAGTTC AATCTGGATT CATATTTAAT TTGAAGGTAG AATGTTTTCA  
 30 701 AATGTTCTCC AGTCATTATG TTAATATTTT TGAGGAGCCT GCAACATGCC  
 751 AGCCACTGTG ATAGAGGCTG GCGGATCCAA GCAAATGGCC AATGAGATCA  
 801 TTGTGAAGGC AGGGGAATGT ATGTGCACAT CTGTTTTGTA ACTGTTTAGA  
 851 TGAATGTCAG TTGTTATTTA TTGAAATGAT TTCACAGTGT GTGGTCAACA  
 901 TTTCTCATGT TGAAACTTTA AGAACTAAAA TGTCTAAAT ATCCCTTGGA  
 35 951 CATTTTATGT CTTTCTTGTA AGGCATACTG CTTTGTTTAA TGGTAGTTTT  
 1001 ACAGTGTTTC TGGCTTAGAA CAAAGGGGCT TAATTATTGA TGTTTTCGGA

-10-

which corresponds to the cDNA sequence and derived amino acid sequence of *gro*/MGSA (melanoma growth stimulatory activity). See Anisowicz et al., Proc. Natl. Acad. Sci. U.S.A. 84, 7188-7192 (1987).

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 18) of:

Met Ala Arg Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu Arg Val Ala Leu  
Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala Ala Gly Ala Pro Lys Ala Thr Glu  
Lys Arg Cys Gln Cys Lys Gln Thr Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val  
Lys Val Lys Ser Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn  
Gly Gln Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Lys Lys Ile Ile Glu Lys Met  
Leu Lys

encoded by the cDNA sequence (SEQ ID NO: 19) of:

1 CTCTCCTCCT CGCACAGCCG CTCGAACCGC CTGCTGAGCC CCATGGCCCCG  
51 CGCCACGCTC TCCGCCGCC CCAGCAATCC CCGGCTCCTG CGGGTGGCGC  
15 101 TGCTGCTCCT GTCCTGGTG GCCGCCAGCC GCGCGCAGC AGGAGCGCCC  
151 CTGGCCACTG AACTGCGCTG CCAGTGCTTG CAGACCCTGC AGGGAATTCA  
201 CCTCAAGAAC ATCCAAAGTG TGAAGGTGAA GTCCCCCGGA CCCCACTGCG  
251 CCCAAACCGA AGTCATAGCC AACTCAAGA ATGGGCAGAA AGCTTGCTCTC  
301 AACCCCGCAT CGCCCATGGT TAAGAAAATC ATCGAAAAGA TGCTGAAAAA  
20 351 TGGCAAATCC AACTGACCAG AAGGAAGGAG GAAGCTTATT GGTGGCTGTT  
401 CCTGAAGGAG GCCCTGCCCT TACAGGAACA GAAGAGGAAA GAGAGACACA  
451 GCTGCAGAGG CCACCTGGAT TCGCCTAAT GTGTTTGAGC ATCACTTAGG  
501 AGAAGTCTTC TATTTATTTA TTTATTTATT TATTTGTTG TTTAGAAGA  
551 TTCTATGTTA ATATTTTATG TGTAATAA GGTATGATT GAATCTACTT  
25 601 GCACACTCTC CCATTATATT TATTGTTTAT TTAGGTCAA ACCCAAGTTA  
651 GTTCAATCCT GATTCATATT TAATTTGAAG ATAGAAGGTT TGCAGATATT  
701 CTCTAGTCAT TTGTTAATAT TTCTTCGTGA TGACATATCA CATGTCAGCC  
751 ACTGTGATAG AGGCTGAGGA ATCCAAGAAA ATGGCCAGTG AGATCAATGT  
801 GACGGCAGGG AAATGTATGT GTGTCTATTT TGTAAGTGA AAGATGAATG  
30 851 TCAGTTGTTA TTTATTGAAA TGATTTTACA GTGTGTGGTC AACATTTCTC  
901 ATGTTGAAGC TTAAAGAACT AAAATGTTCT AAATATCCCT TGGACATTTT  
951 ATGTCTTTCT TGTAAGGCAT ACTGCCTTGT TTAATGTTAA TTATGCAGTG  
1001 TTTCCCTCTG TGTTAGAGCA GAGAGGTTTC GATATTTATT GATGTTTTCA  
1051 CAAAGAACAG GAAAATAAAA TATTTAAAAA T

35 which corresponds to the cDNA sequence and derived amino acid sequence *gro*- $\beta$ /MIP-2 $\alpha$  (macrophage inflammatory protein 2- $\alpha$ ). See Tekamp-Olson et al., J. Exp. Med. 172, 911-919

(1990).

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 20) of:

Met Ala His Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu Arg Val Ala Leu  
 5 Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala Ala Gly Ala Ser Val Val Thr Glu  
 Leu Arg Cys Gln Cys Leu Gln Thr Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val  
 Asn Val Arg Ser Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn  
 Gly Lys Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Gln Lys Ile Ile Glu Lys Ile Leu  
 Asn Lys Gly Ser Thr Asn

10 encoded by the cDNA sequence (SEQ ID NO: 21) of:

1 CTCGCACAGC TTCCCAGCGC GTCTGCTGAG CCCCATGGCC CACGCCACGC  
 51 TCTCCGCCGC CCCAGCAAT CCCCGGCTCC TCGGGGTGGC GCTGCTGCTC  
 101 CTGCTCCTGG TGGCCGCCAG CCGGCGCGCA GCAGGAGCGT CCGTGGTCAC  
 151 TGAAGTGGC TGCCAGTGCT TGCAGACACT GCAGGGAATT CACCTCAAGA  
 15 201 ACATCCAAAG TGTGAATGTA AGGTCCCCCG GACCCCACTG CGCCCAAACC  
 251 GAAGTCATAG CCACACTCAA GAATGGGAAG AAAGCTTGTC TCAACCCCGC  
 301 ATCCCCCATG GTTCAGAAAA TCATCGAAAA GATACTGAAC AAGGGGAGCA  
 351 CCAACTGACA GGAGAGAAGT AAGAAGCTTA TCAGCGTATC ATTGACACTT  
 401 CCTGCAGGGT GGTCCCTGCC CTTACCAGAG CTGAAAATGA AAAAGAGAAC  
 20 451 AGCAGCTTTC TAGGGACAGC TGGAAAGGAC TTAATGTGTT TGAATATTTT  
 501 TTACGAGGGT TCTACTTATT TATGTATTTA TTTTGAAG CTTGTATTTT  
 551 AATATTTTAC ATGCTGTTAT TTAAAGATGT GAGTGTGTTT CATCAAACAT  
 601 AGCTCAGTCC TGATTATTTA ATTGGAATAT GATGGGTTTT AAATGTGTCA  
 651 TTAAACTAAT ATTTAGTGGG AGACCATAAT GTGTCAGCCA CCTTGATAAA  
 25 701 TGACAGGGTG GGGAACTGGA GGGTGGGGGG ATTGAAATGC AAGCAATTAG  
 751 TGGATCACTG TAGGGTAAG GGAATGTATG TACACATCTA TTTTATATAC  
 801 TTTTATTTTA AAAAAAGAAT GTCAGTTGTT ATTTATTCAA ATTATCTCAC  
 851 ATTATGTGTT CAACATTTTT ATGCTGAAGT TTCCCTTAGA CATTATATGT  
 901 CTTGCTTGTA GGGCATAATG CCTTGTTTAA TGTCCATTCT GCAGCGTTTC  
 30 951 TCTTCCCTT GGAAAAGAGA ATTTATCATT ACTGTTAC

which corresponds to the cDNA sequence and derived amino acid sequence *gro-γ*/MIP-2β (macrophage inflammatory protein 2-β). See Tekamp-Olson et al., J. Exp. Med. 172, 911-919 (1990).

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 22) of:

Met Thr Ser Lys Leu Ala Val Ala Leu Leu Ala Ala Phe Leu Ile Ser Ala Ala Leu Cys

-12-

Glu Gly Ala Val Leu Pro Arg Ser Ala Lys Glu Leu Arg Cys Gln Cys Ile Lys Thr Tyr  
 Ser Lys Pro Phe His Pro Lys Phe Ile Lys Glu Leu Arg Val Ile Glu Ser Gly Pro His Cys  
 Ala Asn Thr Glu Ile Ile Val Lys Leu Ser Asp Gly Arg Glu Leu Cys Leu Asp Pro Lys  
 Glu Asn Trp Val Gln Arg Val Val Glu Lys Phe Leu Lys Arg Ala Glu

5 encoded by the cDNA sequence (SEQ ID NO: 23) of:

1 ATGACTTCCA AGCTGGCOGT GGCTCTCTTG GCAGCCTTCC TGATTTCTGC  
 51 AGCTCTGTGT GAAGGTGCAG TTTTGCCAAG GAGTGCTAAA GAACTTAGAT  
 101 GTCAGTGCAT AAAGACATAC TCCAAACCTT TCCACCCCAA ATTTATCAAA  
 151 GAACTGAGAG TGATTGAGAG TGGACCACAC TGCGCCAACA CAGAAATTAT  
 10 201 TGTAAGCTT TCTGATGGAA GAGAGCTCTG TCTGGACCCC AAGGAAAAC  
 251 GGGTGCAGAG GGTGTGGAG AAGTTTTTGA AGAGGGCTGA G

which corresponds to the cDNA sequence and derived amino acid sequence Interleukin-8/NAP-1 (neutrophil activating protein-1). See Kunser et al., *Kidney Int.* 39, 1240-1248 (1991).

In another aspect, the present invention provides a heparanase having the amino acid  
 15 sequence (SEQ ID NO: 24) of:

Ala Gly Pro Ala Ala Val Leu Arg Glu Lys Arg Cys Val Cys Leu Gln Thr Thr Gln  
 Gly Val His Pro Lys Met Ile Ser Asn Leu Gln Val Phe Ala Ile Gly Pro Gln Cys Ser  
 Lys Val Glu Val Val Ala Ser Leu Lys Asn Gly Lys Glu Ile Cys Leu Asp Pro Glu Ala  
 Pro Phe Leu Lys Lys Val Ile Gln Lys Ile Leu Asp Gly Gly Asn Lys Glu Asn

20 encoded by the cDNA sequence (SEQ ID NO: 25) of:

1 GTGTTGCGGG AACTGCGGTG CGTGTGTTTA CAGACCACGC AGGGAGTTCA  
 51 TCCCAAAATG ATCAGTAATC TGCAAGTGTT CGCCATAGGC CCACAGTGCT  
 101 CCAAGGTGGA AGTGGTAGCC TCCCTGAAGA ACGGGAAGGA AATTTGTCTT  
 151 GATCCAGAAG CCCCTTTTCT AAAGAAAGTC ATCCAGAAAA TCCTCGACGG  
 25 201 CGGCAACAAA GAAAAC

which corresponds to the cDNA sequence and derived amino acid sequence of a novel inflammatory peptide (ENA-78) with homology to interleukin 8. See Walz et al., *J. Exp. Med.* 174, 1355-1362 (1991).

In another aspect, the present invention provides a heparanase having the amino acid  
 30 sequence (SEQ ID NO: 26) of:

Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala Leu Cys Asn Gln  
 Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr Ala Cys Cys Phe Ser Tyr Thr Ser  
 Arg Gln Ile Pro Gln Asn Phe Ile Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys  
 Pro Ser Val Ile Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu

35 Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala

encoded by the cDNA sequence (SEQ ID NO: 27) of:

1 GAATTCAAGG CCTGTCCTGG TTTGGTCCCA ATTTACCTTT ATCATCCATA  
51 TTCACCCCCA CTGCTCTGCA GCTCCACTGA AGCACCCCCT CTTTCCTCTG  
101 AGCCACAATG TCACACCCAG GACTCTGCCT CAGCTGGGCC TCCACTGCCC  
151 ACCCATCTAT AGATGCCTAA ATCCCGGGCA GTTATCCAGA CACAATAAA  
5 201 GTTCCATCCC TTCCATGAAG CCTTCCCCAA CCCTCTGGTG GAAGGTCACT  
251 TCTTCCTCAT GGGGTTCTGA GCTTTCATTT CTTTTTCTAO TAAGAGTTTT  
301 ACAATTACCT GTTCATACAC TCTACCTGCC CCCATGAGAC CAGGGGCATC  
351 TCAGAAACAA AGATCATTAA AACCAACTAA ATCTATTTCT CATTATAAAA  
401 TGAGATATGC TGATTGATTG CAAAATAATA AAATAACAAA GTATGGAAAA  
10 451 GAAAAAAAAA AGCATATAAT CTGGCTGAGA AGGTAGAGAC CCTTCCACAC  
501 CACTGAAATT ATGTGTTGAA AAGAATAAGG AAAAACTGC TTCAGTTTGG  
551 CATTATTTAT GTAAGTATAG TATAGGATCC TAAAAATGGT TCAAAGAAAT  
601 GGGAAATCAA GACTTCATTT TGGCAAAGCC ATTGAACAGA AACTGTAGCA  
651 TATTTATCAG TAATTTCTTT CAGATTAAAC AACTGACAAC AACCCACTTT  
15 701 TCAACCAGTG ATGTTGGAAA TGTTTTAAAA CAAAATTAGT TCATAAATTT  
751 GTGGGTTGAC CAAGAAGGTA ATAAAGTCTC ACTAAATAAA ATGAGGAAAA  
801 TTCAGAAAAA GAAAAAATA AGAAAAATAA TCACCCATGG ATCTAAGCAC  
851 TATTCATTCT TTAAGGCATG TATTTCCAAG CCTTTTAATT TTTTCATGCC  
901 TAGAGTTGGC ATGGCATATA TATATCTTTA TACAATTCTT CAAATTTTAT  
20 951 AGAATTTGTA TAATGTTTTA TCTTGCTTTT TTTTAAACCA CTGATGTTAT  
1001 AAGCATATTT ATGCCACTTC ATTACGTTA GAGACTTAAT AATAAAGGAT  
1051 CTTGTGGATA ATTTATCATT CCCTGATAGA GAAAAATTTA GCTTTGCTTA  
1101 TTTTAGAGTT ATAAATGATG CTGGGTCAGG TATCTTTATG TTTGAAGATG  
1151 GCTCCATATT TGGGTTGTTT CCACAGAACT CTTTCCAGAA ATGCTTTTTT  
25 1201 TAGGTTAATG GCTACACATA TTTCTAGGCA CCTGACATAC TGACACCCAC  
1251 CTCTAAAGTA TTTTATGAT CCACAACCTAG CGTTTAACAC AGCGCCCCAG  
1301 TCACTCCGAG ACTAATAAAT AGACAAATGA CTGAAACGTG ACCTCATGCT  
1351 TTCTATTCCT CCAGCTTTCA TTGAGTTCCT TTCCTCTGGG AGGACTGGGG  
1401 GTTGTCTAGC CCTCCACAGC ATCAGCCCAT TGACCCTATC CTTGTGGTTA  
30 1451 TAGCAGCTGA GGAAGCAGAA TTAGCTCT GTGGGAAGGA ATGGGGCTGG  
1501 AGAGTTCATG CATAGACCAA TTCITTTTTT TTTTTTTTTT TGAGATGGAG  
1551 TTTCACTTTT GTTGCCCAGG CTGGAGTGCA ATGGCATGAT CTCAGCTCAC  
1601 CACAGCCCCC ACCTCCTGGG TTCAAGCGAT TCTCCTGCCC TCAGCCTCCC  
1651 GAGTAGCTGG GATTACAGGC ATGTGCCACC ACGCTGACT ACTTTTGTAT  
35 1701 TTTTAGTAGA GATGGAGTTT CTCTTCTTG GTCAGGTTGG TCTCAAACCTC  
1751 CTGACCTCAG GTGATCTGCA GCCTCGGCCT CCAAAGTGTT GGGATTACAG

1801 GTGTGAGCGA CCATGCCTGG CTGCATAGAC CAGTTCCTTAT GAGAAGGGAT  
1851 CAACTAAGAA TAGCCTTGGG TTGACACACA CCCCTCTTCA CACTCACAGG  
1901 AGAAACCCCA TGAAGCTAGA ACCAGTCATG AGTTGAGAGC TGAGAGTTAG  
1951 AGAGTAGCTC AGAGATGCTA TTCTTGATA TCCTGAGCCC CTGTGGTCAC  
5 2001 CAGGGACCCT GAGTTGTGCA AACTCAGCA TGACAGCATC ACTACACTTA  
2051 AAAATTTCCC TCCTCACCCC CAGATTCCAT TTCCCATCC GCCAGGGCTG  
2101 CCTATAAAGA GGAGAGATGG CTTCAGACAT CAGAAGGACG CAGGCAGCAA  
2151 AGAGTAGTCA GTCCCTTCTT GGCTCTGCTG AACTCGAGC CCACATTCCA  
2201 TCACCTGCTC CCAATCATGC AGGTCTCCAC TGCTGCCCTT GCCGTCTCC  
10 2251 TCTGCACCAT GGCTCTCTGC AACCAGGTCC TCTCTGCACC ACGTGAGTCC  
2301 ATGTTGTTGT TGTGGGTATC ACCACTCTCT GGCCATGGTT AGACCACATC  
2351 AGTCTTTTTT TGTGGCGTGA GAGGCCCCGA AGAGAAAAGA AGGAAGTTCT  
2401 TAAAGCGCTG CCAAACACCT TGGTCTTTTT CTTCACAACT TTTATTTTA  
2451 TCTCTAGAAG GGTCTTAGC CCTCCTAGTC TCCAGGTATG AGAATCTAGG  
15 2501 CAGGGGCAGG GGAGTTACAG TCCCTTGTAC AGATAGAAAA ACAGGGTTCA  
2551 AAACGAATCA GTTTGCAAGA GGCAGAATCC AGGGCTGCTT ACTTCCCAGT  
2601 GGGGTCTGTT CTTCACTCTC CAGCTCACCC TAGTCTCCCA GGAGCCCTGT  
2651 CCCTTGGATG TCTTATGAGA GATGTCCAGG GCTTCTCTTG GGCTGGGGTA  
2701 TGACTTCTTG AACCGACAAA ATTCCATGAA GAGAGCTAAG AGAACAGTCC  
20 2751 ATTCAGGTAT CTGGATCACA TAGAGAAACA GAGAACCAC TATGAAGAGT  
2801 CAAGGGGAAA GAGGAATATA GACAGAAACA AAGAGACATT TCTCTGCAAA  
2851 ACCCCCCAAA TGCCTTGCAG TCACTTGGTC TGAGCAAGCC TGCCCTCTC  
2901 AACCACTCAG GGATCAGAAG CTGCCTGGCC TTTTCTCTG AGCTGTGACT  
2951 TGGGCTTATT CTCTCCTTC TCCGCA GTT CTGCTGACAC GCCGACCGCC  
25 3001 TGCTGCTTCA GCTACACCTC CCGACAGATT CCACAGAATT TCATAGCTGA  
3051 CTACTTTGAG ACGAGCAGCC AGTGCTCCAA GCCCAGTGTC ATGTAAGTGC  
3101 CAGTCTTCCT GCTCACCTCT AGGGAGGTAG GGAGTGTCAG GGTGGGGGCA  
3151 GAAACAGGCC AGAAGGCCAT CCTGGAAAGG CCCAGCCTTC AGGAGCCTAT  
3201 CGGGGATACA GGACGCAGGG CACTGAGGTG TGACCTGACT TGGGGCTGGA  
30 3251 GTGAGGTGGG TGTTACAGAG TCAGGAAGGG CTGCCCCAGG CCAGAGGAAA  
3301 GGGACAGGAA GAAGGAGGCA GCAGGAACT CTGAGGGCCC CCTTGCTGG  
3351 AGTCACTGAG AGAAGCTCTC TAGACGGAGA TAGGCAGGGG GCCCTGAGA  
3401 GAGGAGCAGG CCTTGAGCTG CCCAGGACAG AGAGCAGGAT GTCAGGGCCA  
3451 TGGTGGGCCC AGGATTCCCC GGCTGGATTG CCCAGTGCTT AACTCTTCCT  
35 3501 CCCTTCTCCA CAGCTTCCTA ACCAAGAGAG GCCGGCAGGT CTGTGCTGAC  
3551 CCCAGTGAGG AGTGGGTCCA GAAATACGTC AGTGACCTGG AGCTGAGTGC



-15-

3601 CTGAGGGGTC CAGAAGCTTC GAGGCCAGC GACCTCAGTG GGCCAGTGG  
 3651 GGAGGAGCAG GAGCCTGAGC CTTGGGAACA TCGTGTGAC CTCCACAGCT  
 3701 ACCTCTTCTA TGGACTGGTT ATTGCCAAAC AGCCACACTG TGGGACTCTT  
 3751 CTTAACTTAA ATTTAATTT ATTTATACTA TTTAGTTTTT ATAATTTATT  
 5 3801 TTTGATTTCAG CAGTGTGTTT GTGATTGTTT GCTCTG GAG TTCCCCCTGT  
 3851 CCCCTCCACC TTCCCTCACA GTGTGTCTGG TGACAACCGA GTGGCTGTCA  
 3901 TCGGCCTGTG TAGGCAGTCA TGGCACCAA GGCACCAGAC TGACAAATGT  
 3951 GTATCAGATG CTTTTGTTC GGGCTGTGAT CGGCCTGGG AAATAATAAA  
 4001 GATGTTCTTT TAAACGGTAA ACCAGTATTG AGTTTGGTTT TGTTTTTCTG  
 10 4051 GCAAATCAAA ATCACTGGTT AAGAGGAATC ATAGGCAAAG ATTAGGAAGA  
 4101 GGTGAAATGG AGGGAAATTG GGAGAGATGG GGAGCGCTGC GACAGAGTTA  
 4151 TCCACTTCAC AAAATTCTGG AACATTGAAA CTACGAATAT GTTATAACTG  
 4201 AAATCGTAAT ATGCACGCTC TAGGAGAATT AACTACTTGA ATGGCCACX  
 4251 TTAAGCAGAG TATTCTGTAG GGCATATTCA TGATGAATCA AGCTCTTAAT  
 15 4301 AGCAATTATT TACATTGTTG AGGCTTACTC CTCCTACTGA GTGCTTTTAA  
 4351 TACATTGTTT ATTTAATCTT ACCAATGCAA TAGTACAGCT TAGGTACTAT  
 4401 TAATACCTCC ACTTGACAGA AAAGTAACCC AGGGCTCAGA AAGGTTAGAC  
 4451 AACTTGGCTG AGGTTACACA GCACGTAAAC GGTCAATTGT GTTCCAAAAC  
 4501 TGGACTTTTA TTGAACTACA GACTATGCTG TTAACCATTG ACCAAGTTAT  
 20 4551 TTCCCAAAGT ATGACCCGCC TATACTCAA TCTTACCCA TTCTTTAACA  
 4601 GATGATACTT TATCCATTGC AACCATTCC TGTCAGGATT CTGAGTTGAC  
 4651 ATAGAGTGTT TCAGCAGTGA TTATTTAAGC CAATTACATC AGGATCTTTA  
 4701 GGTGTAGACC TGGGAACTGA TATTTTATC AAGCTCATGA GGTGTTCCAT  
 4751 AGCATGTAA TGA CTGAGAG CCACTGTCAA TAGAATTC

25 which corresponds to the cDNA sequence and derived amino acid sequence MIP-1 $\alpha$   
 (macrophage inflammatory protein 1- $\alpha$ ). See Blum et al., DNA Cell Biol. 9, 589-602 (1990).

In another aspect, the present invention provides a heparanase having the amino acid  
 sequence (SEQ ID NO: 28) of:

30 Met Lys Leu Cys Val Thr Val Leu Ser Leu Leu Met Leu Val Ala Ala Phe Cys Ser Pro  
 Ala Leu Ser Ala Pro Met Gly Ser Asp Pro Pro Thr Ala Cys Cys Phe Ser Tyr Thr Ala  
 Arg Lys Leu Pro Arg Asn Phe Val Val Asp Tyr Tyr Glu Thr Ser Ser Leu Cys Ser Gln  
 Pro Ala Val Val Phe Gln Thr Lys Arg Ser Lys Gln Val Cys Ala Asp Pro Ser Glu Ser  
 Trp Val Gln Glu Tyr Val Tyr Asp Leu Glu Leu Asn

encoded by the cDNA sequence (SEQ ID NO: 29) of:

35 1 TTCCCCCCCC CCCCCCCCCC CCGCCCCG GCACAGGACA CAGCTGGGTT  
 51 CTGAAGCTTC TGAGTTCTGC AGCCTCACCT CTGAGAAAAC CTCTTTTCCA

-16-

101 CCAATACCAT GAAGCTCTGC GTGACTGTCC TGTCTCTCCT CATGCTAGTA  
 151 GCTGCCTTCT GCTCTCCAGC GCTCTCAGCA CCAATGGGCT CAGACCCCTCC  
 201 CACCGCCTGC TGCTTTTCTT ACACCGCGAG GAAGCTTCCT CGCAACTTTG  
 251 TGGTAGATTA CTATGAGACC AGCAGCCTCT GCTCCCAGCC AGCTGTGGTA  
 5 301 TTCCAAACCA AAAGAAGCAA GCAAGTCTGT GCTGATCCCA GTGAATCCTG  
 351 GGTCCAGGAG TACGTGTATG ACCTGGAAGT GAACTGAGCT GCTCAGAGAC  
 401 AGGAAGTCTT CAGGGAAGGT CACCTGAGCC CGGATGCTTC TCCATGAGAC  
 451 ACATCTCCTC CATACTCAGG ACTCCTCTCC GCAGTTCCTG TCCCTTCTCT  
 501 TAATTTAATC TTTTTTATGT GCCGTGTTAT TGTATTAGGT GTCATTTCCA  
 10 551 TTATTTATAT TAGTTTAGCC AAAGGATAAG TGTCTATGG GGATGGTCCA  
 601 CTGTCACTGT TTCTCTGCTG TTGCAAATAC ATGGATAACA CATTTGATTC  
 651 TGTGTGTTTT CCATAATAAA ACTTTAAAAT AAAATGCAGA CAGTTA

which corresponds to the cDNA sequence and derived amino acid sequence MIP-1 $\beta$   
 (macrophage inflammatory protein 1- $\beta$ ). See Lipes et al., Proc. Natl. Acad. Sci. U.S.A. 85,  
 15 9704-9708 (1988).

In another aspect, the present invention provides a heparanase having the amino acid  
 sequence (SEQ ID NO: 30) of:

Met Gln Ile Ile Thr Thr Ala Leu Val Cys Leu Leu Leu Ala Gly Met Trp Pro Glu Asp  
 Val Asp Ser Lys Ser Met Gln Val Pro Phe Ser Arg Cys Cys Phe Ser Phe Ala Glu Gln  
 20 Glu Ile Pro Leu Arg Ala Ile Leu Cys Tyr Arg Asn Thr Ser Ser Ile Cys Ser Asn Glu  
 Gly Leu Ile Phe Lys Leu Lys Arg Gly Lys Glu Ala Cys Ala Leu Asp Thr Val Gly Trp  
 Val Gln Arg His Arg Lys Met Leu Arg His Cys Pro Ser Lys Arg Lys

encoded by the cDNA sequence (SEQ ID NO: 31) of:

1 ACCAGGCTCA TCAAAGCTGC TCCAGGAAGG CCCAAGCCAG ACCAGAAGAC  
 25 51 ATGCAGATCA TCACCACAGC CCTGGTGTGC TTGCTGCTAG CTGGGATGTG  
 101 GCCGGAAGAT GTGGACAGCA AGAGCATGCA GGTACCCTTC TCCAGATGTT  
 151 GCTTCTCATT TCGGAGCAA GAGATTCCTC TGAGGGCAAT CCTGTGTTAC  
 201 AGAAATACCA GCTCCATCTG CTCCAATGAG GGCTTAATAT TCAAGCTGAA  
 251 GAGAGGCAAA GAGGCCTGCG CCTTGGACAC AGTTGGATGG GTTCAGAGGC  
 30 301 ACAGAAAAAT GCTGAGGCAC TGCCCGTCAA AAAGAAAATG AGCAGATTTC  
 351 TTTCCATTGT GGGCTCTGGA AACCACATGG CTTACCTGT CCCCAGAACT  
 401 ACCAGCCCTA CACCATTCCT TCTGCCCTGC TTTTGCTAGG TCACAGAGGA  
 451 TCTGCTTGGT CTGATAAGC TATGTTGTTG CACTTTAAAC ATTTAAATTA  
 501 TACAATCATC AACCCCAAC

35 which corresponds to the cDNA sequence and derived amino acid sequence human secreted  
 protein (I-309). See Miller et al., J. Immunol. 143, 2907-2916 (1989).

-17-

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 32) of:

Met Lys Val Ser Ala Ala Leu Leu Cys Leu Leu Leu Ile Ala Ala Thr Phe Ile Pro Gln  
Gly Lys Ala Gln Pro Asp Ala Ile Asn Ala Pro Val Thr Cys Cys Tyr Asn Phe Thr Asn  
5 Arg Lys Ile Ser Val Gln Arg Leu Ala Ser Tyr Arg Arg Ile Thr Ser Ser Lys Cys Pro Lys  
Glu Ala Val Ile Phe Lys Thr Ile Val Ala Lys Glu Ile Ala Ala Pro Lys Gln Lys Trp  
Val Gln Asp Ser Met Asp His Leu Asp Lys Gln Thr Gln Thr Pro Lys Thr

encoded by the cDNA sequence (SEQ ID NO: 33) of:

1 CTAACCCAGA AACATCCAAT TCTCAAAGT AAGCTCGCAC TCTCGCCTCC  
10 51 AGCATGAAAG TCTCTGCCGC CCTTCTGTGC CTGCTGCTCA TAGCAGCCAC  
101 CTTCAATCCC CAAGGGCTCG CTCAGCCAGA TGCAATCAAT GCCCCAGTCA  
151 CCTGCTGTTA TAACTTCACC AATAGGAAGA TCTCAGTGCA GAGGCTCGCG  
201 AGCTATAGAA GAATCACCAG CAGCAAGTGT CCCAAAGAAG CTGTGATCTT  
251 CAAGACCATT GTGGCCAAGG AGATCTGTGC TGACCCCAAG CAGAAGTGGG  
15 301 TTCAGGATTC CATGGACCAC CTGGACAAGC AAACCCAAAC TCCGAAGACT  
351 TGAACACTCA CTCCACAACC CAAGAATCTG CAGCTAACTT ATTTTCCCTT  
401 AGCTTTCCCC AGACACCCTG TTTTATTTTA TTATAATGAA TTTGTTTGT  
451 TGATGTGAAA CATTATGCCT TAAGTAATGT TAATTCTTAT TTAAGTTATT  
501 GATGTTTTAA GTTTATCTTT CATGGTACTA GTGTTTTTTA GATACAGAGA  
20 551 CTTGGGGAAA TTGCTTTTCC TCTTGAACCA CAGTTCTACC CCTGGGATGT  
601 TTTGAGGGTC TTTGCAAGAA TCATTAATAC AAAGAATTTT TTTTAACATT  
651 CCAATGCATT GCTAAAATAT TATTGTGGAA ATGAATATTT TGTAAGTATT  
701 ACACCAAATA AATATATTTT TGTAC

which corresponds to the cDNA sequence and derived amino acid sequence monocyte  
25 chemoattractant protein 1 (MCP-1). See Yoshimura et al., FEBS Lett. 244, 487-493 (1989).

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 34) of:

Met Lys Ala Ser Ala Ala Leu Leu Cys Leu Leu Leu Thr Ala Ala Ala Phe Ser Pro Gln  
Gly Leu Ala Gln Pro Val Gly Ile Asn Thr Ser Thr Thr Cys Cys Tyr Arg Phe Ile Asn  
30 Lys Lys Ile Pro Lys Gln Arg Leu Glu Ser Tyr Arg Arg Thr Thr Ser Ser His Cys Pro  
Arg Glu Ala Val Ile Phe Lys Thr Lys Leu Asp Lys Glu Ile Cys Ala Asp Pro Thr Gln  
Lys Trp Val Gln Asp Phe Met Lys His Leu Asp Lys Lys Thr Gln Thr Pro Lys Leu

encoded by the cDNA sequence (SEQ ID NO: 35) of:

1 AGCAGAGGGG CTGAGACCAA ACCAGAAACC TCCAATTCTC ATGTGGAAGC  
35 51 CCATGCCCTC ACCCTCCAAC ATGAAAGCCT CTGCAGCACT TCTGTGTCTG  
101 CTGCTCACAG CAGCTGCTTT CAGCCCCCAG GGGCTTGCTC AGCCAGTTGG

151 GATTAATACT TCAACTACCT GCTGCTACAG ATTTATCAAT AAGAAAATCC  
 201 CTAAGCAGAG GCTGGAGAGC TACAGAAGGA CCACCAGTAG CCACTGTCCC  
 251 CGGGAAGCTG TAATCTTCAA GACCAAAGT GACAAGGAGA TCTGTGCTGA  
 301 CCCCACACAG AAGTGGGTCC AGGACTTTAT GAAGCACCTG GACAAGAAAA  
 5 351 CCCAACTCC AAAGCTTTGA ACATTCATGA CTGAACTAAA AACAAGCCAT  
 401 GACTTGAGAA ACAAATAATT TGTATACCCT GTCCTTTCTC AGAGTGGTTC  
 451 TGAGATTATT TTAATCTAAT TCTAAGGAAT ATGAGCTTTA TGTAATAATG  
 501 TGAATCATGG TTTTCTTAG TAGATTTTAA AAGTTATTAA TATTTTAATT  
 551 TAATCTTCCA TGGATTTTGG TGGGTTTTGA ACATAAAGCC TTGGATGTAT  
 10 601 ATGTCATCTC AGTGCTGTAA AACTGTGGG ATGCTCCTCC CTCTCTACC  
 651 TCATGGGGGT ATTGTATAAG TCCTTGCAAG AATCAGTGCA AAGATTGTCT  
 701 TTAATTGTTA AGATATGATG TCCCTATGGA AGCATATTGT TATTATATAA  
 751 TTACATATTT GCATATGTAT GACTCCCAA TTTTCACATA AAATAGATTT  
 801 TTGTAAAAAA

15 which corresponds to the cDNA sequence and derived amino acid sequence monocyte  
 chemoattractant protein 3 (MCP-3). See: Structural and Functional Identification of Two  
 Human, Tumor-derived Monocyte Chemotactic Proteins (MCP-2 and MCP-3) Belonging to the  
 Chemokine Family. Jo Van Damme, Paul Proost, Jean-Pierre Lenaerts, and Ghislain  
 Opdenakker. J. Exp. Med. 176: 59-65, 1992.

20 In another aspect, the present invention provides a heparanase having the amino acid  
 sequence (SEQ ID NO: 36) of:

Met Lys Val Ser Ala Ala Arg Leu Ala Val Ile Leu Ile Ala Thr Ala Leu Cys Ala Pro  
 Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro Cys Cys Phe Ala Tyr Ile Ala Arg Pro  
 Leu Pro Arg Ala His Ile Lys Glu Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala  
 25 Val Val Phe Val Thr Arg Lys Asn Arg Gln Val Cys Ala Asn Pro Glu Lys Lys Trp Val  
 Arg Glu Tyr Ile Asn Ser Leu Glu Met Ser

encoded by the cDNA sequence (SEQ ID NO: 37) of:

1 CCTCCGACAG CCTCTCCACA GGTACCATGA AGGTCTCCGC GGCACGCCTC  
 51 GCTGTCATCC TCATTGCTAC TGCCCTCTGC GCTCCTGCAT CTGCCTCCCC  
 101 ATATTCCTCG GACACCACAC CCTGCTGCTT TGCCTACATT GCCCCGCCAC  
 151 TGCCCCGTGC CCACATCAAG GAGTATTTCT ACACCAGTGG CAAGTGCTCC  
 201 AACCAGCAG TCGTCTTTGT CACCCGAAAG AACCGCCAAG TGTGTGCCAA  
 251 CCCAGAGAAG AAATGGGTTC GGGAGTACAT CAACTCTTTG GAGATGAGCT  
 301 AGGATGGAGA GTCCTTGAAC CTGAACTTAC ACAAATTTGC CTGTTTCTGC  
 35 351 TTGCTCTTGT CCTAGCTTGG GAGGCTTCCC CTCCTATCC TACCCACCCC  
 401 GCTCCTTGAA GGGCCCAGAT TCTGACCACG ACGAGCAGCA GTTACAAAAA

451 CCTTCCCCAG GCTGGACGTG GTGGCTCAGC CTTGTAATCC CAGCACTTTG  
 501 GGAGGCCAAG GTGGGTGGAT CACTTGAGGT CAGGAGTTCG AGACAGCCTG  
 551 GCCAACATGA TGAAACCCCA TGTGTACTAA AAATACAAAA AATTAGCCGG  
 601 GCGTGGTAGC GGGCGCCTGT AGTCCCAGCT ACTCGGGAGG CTGAGGCAGG  
 5 651 AGAATGGCGT GAACCCGGGA GCGGAGCTTG CAGTGAGCCG AGATCGCGCC  
 701 ACTGCACTCC AGCCTGGGCG ACAGAGCGAG ACTCCGTCTC AAAAAAAAAA  
 751 AAAAAAAAAA AAAAAATACA AAAATTAGCC GCGTGGTGGC CCACGCCTGT  
 801 AATCCCAGCT ACTCGGGAGG CTAAGGCAGG AAAATTGTTT GAACCCAGGA  
 851 GGTGGAGGCT GCAGTGAGCT GAGATTGTGC CACTTCACTC CAGCCTGGGT  
 10 901 GACAAAGTGA GACTCCGTCA CAACAACAAC AACAAAAAGC TTCCCCAACT  
 951 AAAGCCTAGA AGAGCTTCTG AGGCGCTGCT TTGTCAAAAG GAAGTCTCTA  
 1001 GGTCTGAGC TCTGGCTTTG CTTGGCTTT GCAAGGGCTC TGTGACAAGG  
 1051 AAGGAAGTCA GCATGCCTCT AGAGGCAAGG AAGGGAGGAA CACTGCACTC  
 1101 TTAAGCTTCC GCCGTCTCAA CCCCTCACAG GAGCTTACTG GCAAACATGA  
 15 1151 AAAATCGGGG

which corresponds to the cDNA sequence and derived amino acid sequence Human T cell-specific protein (RANTES). See Schall et al., J. Immunol. 141, 1018-1025 (1988).

In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 38) of:

20 Met Arg Ile Ser Ala Thr Leu Leu Cys Leu Leu Leu Ile Ala Ala Ala Phe Ser Ile Gln Val  
 Trp Ala Gln Pro Asp Gly Pro Asn Ala Ser Thr Cys Cys Tyr Val Lys Lys Gln Lys Ile  
 Pro Lys Arg Asn Leu Lys Ser Tyr Arg Arg Ile Thr Ser Ser Arg Cys Pro Trp Glu Ala  
 Val Ile Phe Lys Thr Lys Lys Gly Met Glu Val Cys Arg Glu Ala His Gln Lys Trp Val  
 Glu Glu Ala Ile Ala Tyr Leu Asp Met Lys Thr Pro Thr Pro Lys Pro

25 encoded by the cDNA sequence (SEQ ID NO: 39) of:

1 ACTGAAGCCA GCTCTCTCAC TCTCTTCTC CACCATGAGG ATCTCTGCCA  
 51 CGCTTCTGTG CCTGCTGCTC ATAGCCGCTG CTTTCAGCAT CCAAGTGTGG  
 101 GCCCAACCAG ATGGGCCCAA TGCATCCACA TGCTGCTATG TCAAGAAACA  
 151 AAAGATCCCC AAGAGGAATC TCAAGAGCTA CAGAAGGATC ACCAGTAGTC  
 30 201 GGTGTCCCTG GGAAGCTGTT ATCTTCAAGA CAAAGAAGGG CATGGAAGTC  
 251 TGTCGTGAAG CCCATCAGAA GTGGGTCGAG GAGGCTATAG CATACTTAGA  
 301 CATGAAAACC CCAACTCCAA AGCCTTGAAG AAATGTGCCT GAACAGAAAC  
 351 CAACCTAGGA GCCAAGAAGC AAAAATTCCT CACCGCTGTT CTTTCTGAGA  
 401 ACTGTTGATG AAATGTGTTG ATCACGGTCC TAAGGGATAG GAGCTGTCTG  
 35 451 TAGGAATGTG AAACAGTCAC GCCTAAGGAA TGGTCTTTAA GTTATTAATA  
 501 TTTTATTTA ATTAGCCATG TACTTTGGTG TGATTTGAAT GTAAAGCTCT

-20-

551 GGAGACCTCA TGTCACCTTA ACATTGTGTT AGCTGCAGAA TTC

which corresponds to the cDNA sequence and derived amino acid sequence human *fic* (growth factor-activated gene). See Heinrich et al., Molecular and Cellular Biology 13: 2020-2030, 1993.

- 5 In another aspect, the present invention provides a heparanase having the amino acid sequence (SEQ ID NO: 40) of:

Asp Ser Val Ser Ile Phe Ile Thr Cys Cys Phe Asn Val Ile Asn Arg Lys Ile Pro Ile Gln  
Arg Leu Glu Ser Tyr Thr Arg Ile Thr Asn Ile Gln Cys Pro Lys Glu Ala Val Ile Phe Lys  
Thr Gly Lys Glu Val Cys Ala Asp Pro Lys Glu Arg Trp Val Arg Asp Ser Met Lys His  
10 Lys Asp Gln Ile Phe Gln Asn Leu Lys Pro

which corresponds to the cDNA sequence and derived amino acid sequence monocyte chemoattractant protein 2 (MCP-2). See VanDamme et al., J. Exp. Med. 176: 59 - 65, 1992.

- The purified heparanase of the present invention, allows for the convenient selection of compounds having anti-heparanase activity (AHA compounds), i.e. inhibitors of heparanase  
15 activity (IHA), by measuring inhibition of heparanase activity. Inhibition of heparanase activity can be measured utilizing *in vivo* radiolabeled heparan sulfate/heparin. This ligand is radiolabeled to high specific activity by intraperitoneal injection of 0.5mCi of S-35 sulfate into C57 mice bearing a 1-2 cm basement membrane tumor (EHS; Engelbreth, Holm, Swarm tumor). The tumor is harvested after 16 hours and the heparan sulfate proteoglycan extracted in 4  
20 volumes of 6M urea, 20mM Tris pH 6.8, protease inhibitors, 0.15M NaCl and 0.5% triton X-100. The urea extract is chromatographed on an anion exchange column and the proteoglycan is eluted in a linear gradient of NaCl. The radiolabeled proteoglycan is exchanged into a solution of 4.0M guanidine-HCl, 20mM Tris pH 7.4 and applied to a size exclusion column. The proteoglycan peak is pooled and exchanged into 0.15mM NaCl and 20mM Tris pH7.4.

- 25 Purified, radiolabeled proteoglycan is coupled to commercially available agarose support. A quantitative assay of heparanase activity is constructed with the radiolabeled ligand in a multi-well format. Briefly, known quantities of recombinant heparanase are added to a multi-well plate containing equal amounts of radiolabeled ligand in each well. Enzyme-ligand interaction proceeds overnight and the ligand-agarose complex is recovered by centrifugation. Radioactivity  
30 in the liquid phase is determined by scintillation counting and is the measure of enzyme activity. Potential enzyme inhibitors can be evaluated by adding the compound to the solution phase or alternatively adding the assay components to multi-well plates containing preweighed amounts of test compound.

- In addition, the purified heparanase of the subject invention can be used for therapeutic  
35 wound healing or can be immobilized onto filters and used to degrade heparin from the blood of patients post-surgery.

Wound treatment can be achieved by administration to an afflicted individual an effective amount of a pharmaceutical composition comprising the purified heparanase in combination with a pharmaceutically acceptable, preferably slow releasing, carrier. See. e.g. PCT/US90/04772, incorporated herein by reference.

5        Immobilization onto filters can be achieved by the methods well known in the art including those disclosed by Langer et al. in *Biomaterials: Inter-facial Phenomenon and Applications*, eds. Cooper et al, pp 493-509, 1982 and those described in U.S. Patent No. 4,373,023, 4,863,611 and 5,211,850 (all incorporated herein by reference).

10        The purified heparanase of the subject invention can be prepared by the method described in procedure A or procedure B, but preferably procedure A.

#### PROCEDURE A

Reverse transcription of the mRNA from activated human leukocyte-derived cells [preferably lymphocytes, neutrophils, platelets, Jurkatt lymphoma cells, Dami cells (Greenberg et al., *Blood* 72:1968-1977, (1988))] is used to prepare the cDNA for the desired heparanase enzyme (preferably SEQ. ID. NO: 1; optionally SEQ. ID. NO: 3, SEQ. ID. NO: 5, SEQ. ID. NO: 7; SEQ. ID. NO: 13, SEQ. ID. NO: 15, SEQ. ID. NO: 17, SEQ. ID. NO: 19, SEQ. ID. NO: 21, SEQ. ID. NO: 23, SEQ. ID. NO: 25, SEQ. ID. NO: 27, SEQ. ID. NO: 29; SEQ. ID. NO: 31, SEQ. ID. NO: 33 or SEQ. ID. NO: 35), employing standard PCR cloning techniques (described in Sambrook et al., in: *Molecular Cloning, A Laboratory Manual*. Second Edition, 20    1989. Cold Spring Harbor Press). The cDNA encoding the heparanase enzyme is cloned into Xba1/BamH1 sites in the commercially available baculovirus vector pVL 1392 (Pharmingen; San Diego, CA). High titer infectious virus is selected for use in infecting sf9 insect cells (Luckow and Summers, *Bio/Technology*. 6,47 1988). Serum-free medium conditioned by infected sf9 cells is collected after 72 hours. This media is the starting material for purification 25    of recombinant heparanase. Serum-free conditioned media is adjusted to contain 20mM Sodium Acetate, pH 5.0, 0.15M NaCl, 1mM reduced glutathione (GSH), 1mM dithiothreitol (DTT) and 10mM beta-octylglucoside. Medium is applied to a column of cation-exchange resin (Pharmacia) and eluted from the column in a linear gradient of NaCl. Fractions containing heparanase are pooled and diluted to a final salt concentration of 0.15M NaCl. To this solution 30    is added 20mM Tris and the pH adjusted to 7.0. The solution is applied to a column of heparin-Sepharose (Pharmacia) and eluted with a linear salt gradient buffered to pH 5.0 with 20mM Sodium Acetate. Heparanase is concentrated to 0.5mg/ml in an Amicon concentrator fitted with a YM-2 membrane and stored at -80 degrees. For optimal activity (greater than 50 units heparanase activity per  $\mu$ g protein) incubation in the presence of transglutaminase, under 35    reducing conditions, in accordance with the procedure in Example 2, Part C.

## PROCEDURE B

This procedure describes the purification to homogeneity of heparanase (SEQ. ID. NO: 1) from human blood cells or cell lines (such as platelets) under reducing conditions which allow for the occurrence of post-translational modifications that increase the specific activity of heparanase and make it suitable for use in the above described screening assay. The cells are treated with a suitable activator (such as, but not limited to, thrombin or histamine) which allows for the release of enzymes and cytokines from the cell. Reducing agents are added to the supernatant from the activated cells. Suitable reducing agents include dithiothreitol (DTT), dithioerythritol (DTE), reduced glutathione (GSH), and  $\beta$ -mercaptoethanol. The reduced, activated supernatant is chromatographed on a column of immobilized heparin or heparan sulfate under reducing conditions at pH 5, using a salt gradient (such as NaCl, KCl, or other salt) to elute the bound proteins. Fractions containing heparanase activity are pooled and exchanged into any buffer appropriate for the pH of 6.8 and containing 0.15 M NaCl, reducing agents, and non-ionic detergent. This is passed over any suitable anion-exchange column (bed volume of 5 ml or less). The unbound material from this column is adjusted to pH 5 with acid, and is loaded onto any suitable cation-exchange column (bed volume of 5 ml or less), equilibrated in a suitable pH 5 buffer containing 0.15 M NaCl, reducing agents, and non-ionic detergents. The bound protein is eluted from the column with a salt gradient, and the fractions containing heparanase activity are pooled and size fractionated to below 30,000 daltons with 30 K-cut-off membranes. The protein below 30,000 daltons is concentrated by either heparin-sepharose chromatography or by centrifugation through 5 K-cut-off membranes.

The present invention is seen more fully by the examples set forth below.

Example 1: Use of Heparanase as a screen for AHA compounds.

1. Heparan sulfate, metabolically labeled (S-35) to a high-specific activity- as described above for the EHS tumor, prepared by papain digestion of chromatographically purified heparan sulfate proteoglycan is coupled to cyanogen bromide activated Sepharose-6B (Pharmacia) according to manufacturer's instructions.

2.  $^{35}$ S-Heparan sulfate-Sepharose 6B is resuspended in: 0.15 M NaCl, 0.03% human serum albumin, 10  $\mu$ M  $MgCl_2$ , 10  $\mu$ M  $CaCl_2$ , antiproteolytic agents (1  $\mu$ g/ml leupeptin, 2  $\mu$ g/ml antipain, 10  $\mu$ g/ml benzamidin, 10 units/ml aprotinin, 1  $\mu$ g/ml chymostatin, and 1  $\mu$ g/ml pepstatin), and 0.05 M Na acetate, pH 5.6 and 5,000 cpm, in a total volume of 200  $\mu$ l, are aliquoted into each well of a 96 well plate. To each well is added 5 units of activated heparanase and the digestion allowed to proceed overnight at 37 degrees.

3. Separation of digested product is accomplished by centrifugation of the 96 well plate. The supernatant, containing cleaved heparan sulfate is decanted and quantitated by scintillation counting.



4. Inhibitors of heparanase activity can be introduced into the liquid-phase of the assay.

5. A potential inhibitor of heparanase activity would be identified by its ability to reduce the amount of radiolabeled heparan sulfate released into the supernatant by 50% at a concentration of 1  $\mu$ M or less.

Example 2: The preparation of heparanase under reducing conditions as outlined in Procedure B.

Part A:

- Platelet-rich plasma ( $10^9$  platelets/ml; 1800 ml) is obtained from healthy, informed volunteers by plasmapheresis. The plasma is removed from the platelets by centrifugation (Heldin, *et al.*, *Exp. Cell Res.* 109: 429-437, 1977). Platelets suspended in phosphate buffered saline (PBS; 0.1 original volume) are then stimulated with 1 U/ml thrombin for 5 min at 37°C. This concentration of thrombin was reported to release 100% of the heparanase activity from platelets (Oldberg, *et al.*, *Biochemistry* 19: 5755-5762, 1980). Following activation, the thrombin is inactivated by the addition of 100 mM phenylmethylsulfonylfluoride (PMSF), and the platelets are centrifuged at 2000 x g for 30 min at 4°C. The supernatant is stored at -80°C until used for the chromatographic purification of heparanase (Part B).

Part B: Chromatographic purification of heparanase.

1. Heparin-Sepharose Chromatography. Activated platelet supernatants are pooled and adjusted to contain 1 mM GSH and 1 mM DTT. This pool is loaded (2.5 ml/min) onto a column of heparin-sepharose (2.6 x 7.5 cm, 40 ml) equilibrated in 1 mM GSH, 1 mM DTT, 150 mM NaCl, 10 mM NaPO<sub>4</sub>, pH 7.4. After loading the sample, the column is washed with 200 ml of 0.15 M NaCl, 1 mM GSH, 1 mM DTT, 10 mM Na acetate, pH 5, followed by 60 ml of 0.35 M NaCl, 1 mM DTT, 1 mM GSH, 10 mM Na acetate, pH 5. The column is then eluted with a 160 ml linear gradient between 0.35 M NaCl and 1.5 M NaCl in the same buffer. Aliquots of each fraction are used for determination of heparanase activity by the "Purification Assay" described later.
2. Anion-exchange chromatography (For example, DEAE-Sephacel, Pharmacia). The 0.9 M - 1.15 M NaCl fractions from the heparin-sepharose column are concentrated using a stirred cell fitted with a PM-10 membrane, and the buffer is exchanged to 0.15 M NaCl, 1 mM DTT, 1 mM GSH, 10 mM  $\beta$ -octylglucoside, 10 mM sodium phosphate, pH 6.8 (8 ml). This sample is loaded onto a 5 ml column of DEAE-Sephacel equilibrated in the same buffer. After loading, the column is washed to baseline absorbance (280 nm) with the equilibration buffer. The flow-through and wash with equilibration buffer are collected as one pool. The column is then eluted with 10 ml of 0.15 M NaCl, 10 mM  $\beta$ -octylglucoside, 1 mM GSH, 1 mM DTT, 10 mM Na acetate, pH 5, followed by 10 ml of 1.5 M NaCl, 10 mM  $\beta$ -octylglucoside, 1 mM GSH, 1 mM

DTT, 10 mM Na acetate, pH 5. Aliquots of each pool are used for determination of heparanase activity by the "Purification Assay".

3. Cation Exchange. The unbound sample from the DEAE-Sephacel column is adjusted to pH 5 with glacial acetic acid and loaded onto a cation exchange column (Poros HS/F, 4.6 mm x 50 mm; PerSeptive Biosystems), pre-equilibrated with 0.15 M NaCl, 1 mM DTT, 1 mM GSH, 10 mM  $\beta$ -octylglucoside, 10 mM Na acetate, pH 5. The flow rate is 3.0 ml/min and 3 ml fractions are collected. After washing the column with 35 ml of equilibration buffer, the column is developed with a 55 ml linear gradient between 0.15 M and 1.5 M NaCl in the same buffer. 10  $\mu$ l aliquots of the fractions are used for determination of heparanase activity by the "Purification Assay".

4. Size fractionation to < 30 kD and concentration on immobilized heparin (Hi-trap heparin-sepharose, Pharmacia). The activity from the Poros HS/F column is size fractionated by centrifuging through 30,000 molecular weight cut-off filters (Millipore ultrafree-MC 30,000 NMWL filter units). The < 30 kD pool is diluted to contain 0.15 M NaCl, and is loaded onto a 1 ml Hi-trap heparin column, pre-equilibrated with 0.15 M NaCl, 1 mM DTT, 1 mM GSH, 10 mM Na acetate, pH 5. The column is eluted with 1.2 M NaCl in the same buffer and the single eluted peak contains the heparanase activity.

Part C: Properties of the purified heparanase.

The final yield of heparanase protein from 1850 ml platelet-rich plasma was 2.7 mg.

- 20 Protein concentration was determined by the method of Lowry (*J. Biol. Chem.* 193: 265-275, 1951), or if more precise determinations were required, by amino acid analysis on an amino acid analyzer (Beckman 6300). The overall recovery of activity was 8%, with a 4150-fold purification. The preparation was judged to be homogeneous by the presence of a single band of 9000 daltons on an 18% silver-stained SDS-polyacrylamide gel, run according to the method of Laemmli (*Nature* 227: 680-685, 1970).

The pH optimum of the purified heparanase was determined by conducting the "Purification assay" activity between pH 3.5 and 8.0, using a citrate buffer (pH 3.5 - 6.0), citrate-phosphate buffer (pH 6.5 - 7.0), and phosphate buffer (pH 7.5 - 8.). Heparanase was active between pH 5.0 and 8.0, with the optimum pH at 5.8.

- 30 N-terminal amino acid sequencing of heparanase produced by this procedure was performed using a gas/liquid phase Protein Sequencer (Applied Biosystems Inc. Model 470). Phenylthiodantoin amino acids were resolved and quantitated by an on-line HPLC system (Model 120, Applied Biosystems Inc.) with data analysis on a Nelson Analytical System. N-terminal amino acid sequences of the heparanase produced in this example were 85 %
- 35 SEQ. ID. NO: 9 (namely:

-25-

Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala  
 1 5 10 15

Glu Leu Arg),

which is identical to CTAP-III, and 15% SEQ. ID. NO: 10 (namely:

5 Ser Ser Thr Lys Gly Gln Thr Lys Arg Asn Leu Ala Lys Gly Lys Glu),  
 1 5 10 15

which is the precursor form, platelet basic protein. Interestingly, the N-terminal sequence of commercial  $\beta$ -thromboglobulin (namely, Calbiochem (Cat. # 605165), Celsus Laboratories (Cat. # 41705), and Haematologic Technologies (Cat. # HBTG-0210), which has low levels of

10 heparanase activity, was 100% SEQ. ID. NO: 11 (namely:

Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu),  
 1 5 10 15

indicating that the commercial preparation is actually CTAP-III and not  $\beta$ -thromboglobulin.

Chromatofocusing of the heparanase produced by this procedure results in two peaks of  
 15 differing isoelectric points. To perform the chromatofocusing, heparanase is dissolved in 0.025 M imidazole, pH 7.3. The sample is loaded onto a 0.5 x 20 cm column of Polybuffer Exchanger 94 (Pharmacia), equilibrated with 0.025 M imidazole, pH 7.3. Immediately after sample loading, Polybuffer 74 (Pharmacia; 1:8, pH 4) is pumped onto the column at 0.5 ml/min. 2 ml fractions are collected, and the pH of each fraction is determined by a narrow range (pH 4-  
 20 7) pH paper. Aliquots of each fraction are used to determine heparanase activity by the "Purification Assay." All of the activity is associated with an absorbance (280 nm) peak that eluted at pH 4.8 to 5.1, representing approximately 10% of total protein, while 90% of the protein is eluted at pH 7.3 and is inactive. Aliquots of each protein peak are separated from the ampholytes by  $C_4$  reverse phase chromatography. The peak that eluted from the chromatofocusing column at pH 7.3 has N-terminal sequences for platelet basic protein and the processed form, CTAP-III. The peak that is eluted from the chromatofocusing column at pH 4.8 - 5.1 also contains the sequences of platelet basic protein and the processed form, CTAP-III. All of the platelet basic protein processed forms have pI's that are calculated and reported to be greater than 7.6. Thus, the heparanase activity resides in the platelet basic protein and/or the processed  
 25 form, CTAP-III that is modified such that the pI is lowered to 4.8 - 5.1.

Heparanase obtained after chromatofocusing exhibited a specific activity of 80 units/ $\mu$ g protein, using the "Purification Assay." This represents a 1000-fold increase in the specific activity compared to the commercial protein ( $\beta$ -thromboglobulin; 0.075 units/ $\mu$ g protein).

The modification that may be responsible for the lower isoelectric point of active  
 35 heparanase is ADP-ribosylation. ADP-ribosylation (Adenine diphosphate-ribosylation) is a post-translational modification of proteins or DNA in which the ADP-ribose group of NAD (Nicotinamide adenine dinucleotide) is enzymatically transferred to proteins or DNA. Since this

modification adds two negatively charged phosphate groups to a molecule, it would result in a lower isoelectric point. Activated platelet supernatants were incubated in 1 mM DTT, 2 mM  $\text{MgCl}_2$ , 100 mM HEPES, pH 7.4, and 0.5  $\mu\text{M}$  [ $^{32}\text{P}$ ]NAD (Specific activity = 1000 Ci/mmol). The labeled proteins were separated by SDS-polyacrylamide gel electrophoresis on an 18% gel, transferred to PVDF (polyvinylidene difluoride) membrane, and exposed to X-ray film. The autoradiogram demonstrates the incorporation of [ $^{32}\text{P}$ ] into a protein of 8000 daltons. The PVDF membrane was immunoblotted with the anti-Peptide C antisera (1:1500 in PBS containing 5% dry milk, 0.05% Tween-20, 0.15 M NaCl, 20 mM Tris, pH 7.4, 2 hours at room temperature, followed by incubation with peroxidase-labeled goat anti-chicken IgG (1:500 in above buffer, 1 hour room temperature), and reacted with a peroxidase substrate. The immunoblot revealed that the 8000 dalton that was labeled with [ $^{32}\text{P}$ ] was CTAP-III/heparanase. The addition of 200  $\mu\text{M}$  sodium nitroprusside, a spontaneous releaser of nitric oxide, to the ADP-ribosylation reaction resulted in 5-fold more incorporation of [ $^{32}\text{P}$ ] label into CTAP-III/heparanase, suggesting that this modification can be regulated in vivo by nitric oxide.

Finally, in an analogous manner to that of glyceraldehyde-3-phosphate dehydrogenase, another platelet ADP-ribosylated glycolytic enzyme (Zhang and Snyder, *Proc. Natl. Acad. Sci. USA* 89: 9382-9385), it was determined that CTAP-III/heparanase has an auto-ADP-ribosylation activity, since the [ $^{32}\text{P}$ ]-ADP-ribosylation of CTAP-III/heparanase occurs in reactions where the only protein present is commercial CTAP-III or purified heparanase. Other chemokine family members tested, which includes IP-1-, IL-8,  $\text{gro-}\alpha$ , and MCAF, also have auto-ADP-ribosylation activity.

It is contemplated that the high specific activity of CTAP-III/heparanase is a consequence of ADP-ribosylation of the enzyme in the presence of nitric oxide. It is further contemplated that the action of transglutaminase on the ADP-ribosylated enzyme will lead to further increase in the specific activity.

An amino acid composition of the heparanase produced in Example 2 gave the expected amino acid composition for CTAP-III and N-terminal sequencing revealed sequences for platelet basic protein and the processed form, CTAP-III, confirming that the heparanase activity is contained in this set of processed proteins and is not due to a minor contaminant. The presence of heparanase activity in three commercial sources of  $\beta$ -thromboglobulin also confirms this conclusion. In addition, polyclonal antibodies to  $\beta$ -thromboglobulin were found to precipitate 30 - 70% of the heparanase activity in three separate experiments, providing additional confirmation.

The activation of heparanase with transglutaminase (prepared in accordance with Example 2, Part B) results in a substantial (about 13-fold) increase in the specific activity of the enzyme. The heparanase (2  $\mu\text{l}$  at 56 nM) obtained by Example 2, Part B is treated with either

-27-

transglutaminase from guinea pig liver (4 mU; Sigma) or with Factor XIII (1 µg; Celsus Laboratories, Inc.), the blood coagulation factor that is activated by treatment with 5 units of thrombin at 37 degrees for 30 minutes. Heparanase is activated by incubation of either 2mU liver transglutaminase or 5 units of activated Factor XIII in the presence of 0.1M NaAcetate buffer at pH 6.0 containing 1mM reduced glutathione and 1mM CaCl for 35 minutes at 37 degrees. Treatment of heparanase with either type of transglutaminase results in a substantial increase in the specific activity of the heparanase.

The high degree of sequence identity between CTAP-III and Interleukin-8, a CXC chemokine family member, assures that an essentially identical folding pattern will be shared by the two proteins. Since the 3-dimensional structure of Interleukin-8 is known (Clare, *et al.*, *Biochemistry* 29: 1689-1696, 1990; Baldwin, *et al.*, *J. Biol. Chem.* 265: 6851-6853), one can model the same for CTAP-III. Such a model can serve to direct research into rationally designed IHA and to help explain the action of transglutaminase in activating the CTAP-III.

#### Part D: Purification Assay for Heparanase Activity

Heparanase activity from platelets or column fractions is detected by its ability to digest the  $\geq 70$  kD  $^{35}\text{S}$ -HSPG to produce lower molecular weight products. Each digest contains 10 µl sample,  $^{35}\text{S}$ -HSPG (2000 cpm), 0.15 M NaCl, 0.03% human serum albumin, 10 µM  $\text{MgCl}_2$ , 10 µM  $\text{CaCl}_2$ , antiproteolytic agents (1 µg/ml leupeptin, 2 µg/ml antipain, 10 µg/ml benzamidine, 10 units/ml aprotinin, 1 µg/ml chymostatin, and 1 µg/ml pepstatin), and 0.05 M Na acetate, pH 5.6 in a total volume of 300 µl. Digests are carried out for 3 to 21 h. The presence of lower molecular weight radiolabeled products is detected by centrifugation through 30,000 MW-cutoff filters. The digests containing 2000 cpm of  $^{35}\text{S}$ -HSPG ( $> 70$  K) are centrifuged through 30,000 molecular weight cut-off filters (Millipore Ultrafree-MC 30,000 NMWL filter units).  $^{35}\text{S}$ -HSPG degradation is evident by the presence of radioactivity in the filtrate that passed through the 30 K membrane; this heparanase activity is expressed as the % of total cpm  $< 30$  K for a given digest. Analysis of heparan sulfate degradation by this method is quick and reproducible. 1 unit of heparanase activity is defined as 1% cpm  $< 30$  K per h. For pH optimum determination, the 0.1 M Na acetate buffer is replaced by 50 mM citrate, citrate-phosphate, or phosphate buffer at varying pH's. For samples from chromatographic steps performed under reducing conditions (1 mM GSH, 1 mM DTT), the concentration of a thiol oxidant (diamide) needed for optimum activity is determined. This concentration (100 µM diamide) is added to all assay tubes when reduced samples are assayed.

#### Preparation of $^{35}\text{S}$ -HSPG ( $>70$ K) for use in the "Purification Assay."

$^{35}\text{S}$ -HSPG ( $>70$  K) is prepared from mice bearing a basement membrane tumor that overproduces HSPG (EHS tumor), using modifications of the method of Ledbetter, *et al.*, 1987. Briefly, the radiolabeled HSPG was prepared by injecting C57BL mice bearing the EHS tumor

(Orkin, et.al., 1977) with sodium [ $^{35}\text{S}$ ]sulfate (0.5 mCi/mouse) 18 h before harvesting the tumor. The HSPG is extracted from the weighed tumor with 6 volumes (w/v) of Buffer A (3.4 M NaCl, 0.1 M 6-aminohexanoic acid, 0.04 M EDTA, 0.008 M *N*-ethylmaleimide, 0.002 M PMSF, and 0.05 M Tris-HCl, pH 6.8), by homogenization with a Polytron for 30 s, followed by stirring at 4°C for 1 h. Insoluble material is collected by centrifugation (12,000 x g for 10 min), and the supernatant is discarded. The insoluble residue is reextracted with 2 volumes (original tumor weight) of Buffer A for 30 min with stirring at 4°C. Insoluble material is again collected by centrifugation, and the supernatant fraction is discarded. The insoluble material is then suspended in 6 volumes of Buffer B (6 M urea, 0.1 M 6-aminohexanoic acid, 0.04 M ethylenediaminetetraacetic acid (EDTA), 0.002 M PMSF, and 0.05 M Tris-HCl, pH 6.8), homogenized with an electric homogenizer (Polytron) for 30 s, and stirred for 2 h at 4°C. The mixture is centrifuged to remove insoluble material, and the supernatant is retained. The insoluble material is reextracted with 2 volumes of Buffer B. The mixture is centrifuged, and the supernatant is combined with the previous supernatant.

$^{35}\text{S}$ -HSPG is isolated from the Buffer B supernatant by sequential chromatography on anion exchange and gel filtration columns. The Buffer B supernatant is dialyzed overnight against 10 volumes of 6 M urea, 0.15 M NaCl, 0.05 M Tris-HCl, pH 6.8, and is adjusted to contain 0.5% non-ionic detergent (Triton X-100). This supernatant (from 11 g tumor) is chromatographed on a 30 ml column of anion exchange resin (DEAE-Sephacel) equilibrated with 6 M urea, 0.15 M NaCl, 0.05% Triton X-100, 0.05 M Tris-HCl, pH 6.8. After loading the supernatant and washing with the equilibration buffer, the column is developed with a 250 ml linear gradient between 0.15 M NaCl and 1.15 M NaCl (Flow = 2.0 ml/min). Fractions are sampled for radioactivity, and those containing the  $^{35}\text{SO}_4$  label that elutes from the DEAE-Sephacel between 0.4 M and 0.8 M NaCl are pooled. The proteoglycan is precipitated by the addition of 4 volumes of 100% EtOH at -20°C overnight. The precipitate is collected by centrifugation and is solubilized in 1 ml of Buffer C (4 M Gu-HCl, 20 mM Tris-HCl, pH 7.2). This solubilized pellet is used for chromatography on a calibrated gel filtration column (1.0 x 50 cm column of Superose 6; Pharmacia) equilibrated in Buffer C (Flow = 0.5 ml/min). Fractions are sampled for radioactivity, and those containing the  $^{35}\text{SO}_4$  label that elutes with a molecular weight  $\geq 70$  kD were pooled. The proteoglycan is precipitated with 100% EtOH as described above. The pellet is dissolved in 3 ml PBS, and dialyzed against 3 x 100 volumes of PBS. Each preparation of  $^{35}\text{S}$ -HSPG is confirmed to be  $\geq 98\%$  heparan sulfate by susceptibility to low pH nitrous acid degradation (Shiveley and Conrad, *Biochemistry* 15: 3932-3942, 1976).

Example 3: Preparation of cDNA encoding Heparanase.

Media is removed from cultured HEL (HEL 92.1.7; Human erythroleukemia; ATCC No. TIB 180) cells stimulated with 10nM phorbol 12-myristate 13-acetate (Sigma Chemical Co., St.

Louis, MO) and the cells scraped from the dish and pelleted by centrifugation. The pellet is extracted with 200 $\mu$ l of TRI reagent (Molecular Research Center Inc. Cincinnati, OH) and the total cellular RNA is prepared according to the manufacturer's instructions. To prepare first strand synthesis the reverse transcriptase reaction was performed with 10 $\mu$ l of total cellular RNA

5 in the presence of 4 $\mu$ l of 5x transcriptase buffer (Bethesda Research Laboratories, Gaithersburg, MD), 1 $\mu$ l 0.2mM DTT, 4 $\mu$ l random hexanucleotides (Amersham Corp. Arlington Heights, ILL), and 1 $\mu$ l 10mM dNTP (BRL). This solution is heated to 95 degrees C for 5 minutes and then placed on ice. To this is added 1 $\mu$ l RNasin and 1 $\mu$ l reverse transcriptase (M-MLV), (Promega, Madison WI). This is incubated at 37 degrees for 60 minutes and then placed on ice. The

10 polymerase chain reaction is carried out as follows. To 3 $\mu$ l of the first strand (above) is added 1 $\mu$ l of each Primer (see below), 77 $\mu$ l of water 10 $\mu$ l 10x PCR buffer (Perkin Elmer Cetus, Norwalk CT) and 2 $\mu$ l each dNTP. This solution is heat denatured at 95 degrees C and 1 $\mu$ l Amplitaq DNA polymerase (Perkin Elmer Cetus) is added. Hybridization temperature begins at 72 degrees and is lowered by one degree per cycle until reaching 55 degrees. Each

15 hybridization step is followed with a constant elongation temperature of 72 degrees. Upon completion the solution is left at 0 degrees until storage at -20 degrees. The products of the PCR reaction are electrophoresed on 3% NuSieve, 1% agarose gels and bands of expected size are excised and purified by standard procedures.

Primers:

20 Platelet Basic Protein: TGG ACT AGT ATG TCC TCC ACC AAA GGA CAA ACT AA  
 CTAP III: TGG ACT AGT ATG AAC TTG GCG AAA GAG GA  
 B-thromboglobulin: TGG ACT AGT ATG GGC AAA GAG GAA AGT CTA GAC AG  
 NAP-2: TGG ACT AGT ATG GAA CTC CGC TGC ATG TGT ATA AA

Example 4: Preparation of cDNA encoding Heparanase.

25 Media is removed from cultured leukocyte-derived cells [e.g., lymphocytes, neutrophils, platelets, Jurkatt lymphoma cells, Dami cells (Greenberg et al., Blood 72:1968-1977, (1988)], stimulated with Concanavalin A or phorbol 12-myristate 13 acetate (Sigma Chemical Co., St. Louis, MO) and the cells scraped from the dish and pelleted by centrifugation. The pellet is extracted with 200 $\mu$ l of TRI reagent (Molecular Research Center Inc. Cincinnati, OH) and the

30 total cellular RNA is prepared according to the manufacturer's instructions. To prepare first strand synthesis the reverse transcriptase reaction was performed with 10 $\mu$ l of total cellular RNA in the presence of 4 $\mu$ l of 5x transcriptase buffer (Bethesda Research Laboratories, Gaithersburg, MD), 1 $\mu$ l 0.2mM DTT, 4 $\mu$ l random hexanucleotides (Amersham Corp. Arlington Heights, ILL), and 1 $\mu$ l 10mM dNTP (BRL). This solution is heated to 95 degrees C for 5 minutes and then

35 placed on ice. To this is added 1 $\mu$ l RNasin and 1 $\mu$ l reverse transcriptase (M-MLV), (Promega,

-30-

- Madison WI). This is incubated at 37 degrees for 60 minutes and then placed on ice. The polymerase chain reaction is carried out as follows. To 3 $\mu$ l of the first strand (above) is added 1 $\mu$ l of each Primer (see below), 77 $\mu$ l of water 10 $\mu$ l 10x PCR buffer (Perkin Elmer Cetus, Norwalk CT) and 2 $\mu$ l each dNTP. This solution is heat denatured at 95 degrees C and 1 $\mu$ l
- 5 Amplitaq DNA polymerase (Perkin Elmer Cetus) is added. Hybridization temperature begins at 72 degrees and is lowered by one degree per cycle until reaching 55 degrees. Each hybridization step is followed with a constant elongation temperature of 72 degrees. Upon completion the solution is left at 0 degrees until storage at -20 degrees. The products of the PCR reaction are electrophoresed on 3% NuSieve, 1% agarose gels and bands of expected size
- 10 are excised and purified by standard procedures.

Primers:

- Platelet Basic Protein: TGG ACT AGT ATG TCC TCC ACC AAA GGA CAA ACT AA  
 CTAP III: TGG ACT AGT ATG AAC TTG GCG AAA GAG GA  
 B-thromboglobulin: TGG ACT AGT ATG GGC AAA GAG GAA AGT CTA GAC AG  
 15 NAP-2: TGG ACT AGT ATG GAA CTC CGC TGC ATG TGT ATA AA

All temperatures expressed throughout the subject specification are in degrees Centigrade.

- The cDNA encoding heparanase is preferably cloned into a vector designed for expression in eukaryotic cells, rather than into a vector designed for expression in prokaryotic cells (e.g. *E. coli*). Eukaryotic cells are preferred for expression of genes obtained from higher
- 20 eukaryotes because the signals for synthesis, processing, and secretion of these proteins are usually recognized, whereas this is often not true for prokaryotic hosts (Ausubel, *et al.*, ed., in Short Protocols in Molecular Biology, 2nd edition, John Wiley & Sons, publishers, pg.16-49, 1992.). Eukaryotic hosts may include, but are not limited to, the following: insect cells, African green monkey kidney cells (COS cells), Chinese hamster ovary cells (CHO cells), and
- 25 Murine 3T3 fibroblasts.

Experiments demonstrating that a synthetic peptide of CTAP-III/NAP-2 or antisera raised against a synthetic peptide of CTAP-III/NAP-2 inhibit the heparanase activity of CTAP-III/NAP-2 suggest that the amino acids participating in enzymatic catalysis are contained in a C-terminal region of the enzyme.

- 30 Peptide Synthesis: A C-terminal peptide contained within the sequences known for CTAP-III (SEQ ID NO: 1), Platelet Basic Protein (SEQ ID NO: 3),  $\beta$ -thromboglobulin (SEQ ID NO: 5), and NAP-2 (SEQ ID NO: 7), and an N-terminal peptide contained within the sequences known for CTAP-III (SEQ ID NO: 1) and Platelet Basic Protein (SEQ ID NO: 3) were synthesized according to standard procedures. The N-terminal peptide has the following sequence (SEQ ID
- 35 NO: 41: Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Cys, in which the final Cys residue was added to regions of known sequence (SEQ ID NOS: 1,3) for the purpose of



conjugation to a carrier protein. The C-terminal peptide has the following sequence (SEQ ID NO: 42): Cys Asn Gln Val Glu Val Ile Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro Arg Ile Lys Lys Ile Val Gln Lys Lys encoded by the cDNA sequence (SEQ ID NO: 43) of TGCAACCAAG TCGAAGTGAT AGCCACACTG AAGGATGGGA  
 5 GGAAATCTG CCTGGACCCA GATGCTCCCA GAATCAAGAA AATTGTACAG  
 AAAAAA. These peptides (SEQ ID NOS: 41 and 42) were produced by stepwise solid phase peptide synthesis on an Applied Biosystems 430A Peptide Synthesizer. The 9-fluoroenylmethoxycarbonyl (Fmoc) group was used as the N<sup>α</sup> amino protecting group, and temporary side-chain protectin groups were as follows: Arg (Pmc), Asn (Trt), Asp (OtBu), Gln  
 10 (Trt), Glu (OtBu), His (Trt), Lys (Boc), Ser (tBu), Thr (tBu). Each residue was single coupled using a HBTU/NMP protocol and capped with acetic anhydride before the next synthesis cycle. After removal of the N-terminal Fmoc group, temporary side-chain protecting groups were removed and the peptide cleaved from the resin by treatment with 95% TFA/5% scavengers (ethyl methyl sulfide/anisole/1,2-ethanedithiol, 1:3:1) for two hours at room temperature. The  
 15 crude peptides were precipitated from the cleavage solution with cold diethyl ether. The precipitated peptide was collected on a sintered glass funnel, washed with diethyl ether, dissolved in dilute acetic acid, evaporated to dryness under reduced pressure, and the residue was redissolved and lyophilized from glacial acetic acid. The crude peptides were purified by preparative reverse phase chromatography on a Phenomenex C-18 column (22.5 x 250 mm)  
 20 using a water/acetonitrile gradient, each phase containing 0.1% trifluoroacetic acid (TFA). Clean fractions, as determined by analytical HPLC, were pooled, the acetonitrile was evaporated under reduced pressure, and the aqueous solution was lyophilized. The purified peptides were characterized by time of flight or FAB mass spectroscopy.

Further, SEQ ID NO: 42 can be produced by recombinant DNA methodology as stated  
 25 in Procedure A (page 21).

**Antisera Production:** The synthetic peptides of CTAP-III/NAP-2 were conjugated to keyhole limpet hemocyanin utilizing a maleimide-activated carrier protein (Pierce Chemical Co. #77107). 300 µg of conjugated peptides were injected into chickens using Freund's complete adjuvant. The antisera were collected 5 weeks after initial immunization. Specific recognition by the  
 30 antisera of commercial CTAP-III (2.5 µg, (Celsus Laboratories Inc., Cincinnati, Ohio; Cat #. 41705), isolated heparanase (1.5 µg), and 10 µl of the platelet supernatant used for purification was achieved by separating the proteins on a reducing 18% polyacrylamide gel (Novex), transferring to nitrocellulose, and incubating with the pre-immune or antisera (1:1500), followed by incubation with a peroxidase labeled goat anti-chicken IgG (1:500; Kierkegaard and Perry) in  
 35 the presence of PBS containing 5% dry milk and 0.05% Tween-20. The pre-immune sera did not recognize 7 - 10 kD proteins in the commercial CTAP-III, isolated heparanase, or platelet

supernatants.

**Inhibition of heparanase activity by the C-terminal synthetic peptide (SEQ ID NO: 42) or antisera:**

- For experiments designed to determine whether the peptide antisera was able to inhibit heparanase activity, the pre-immune and antisera were exchanged into 0.15M NaCl, 0.01M sodium phosphate buffer, pH 7.4 (PBS) using a 100 kD cut-off membrane in order to remove low molecular weight chicken heparanase normally present in the serum. Aliquots of isolated heparanase (15 ng) were pre-incubated for 30 min with 2  $\mu$ l of either pre-immune or anti-CTAP-III antisera before adding the  $^{35}$ S-HSPG to determine heparanase activity. In the presence of the pre-immune sera, the isolated protein had  $14.3 \pm 0.1$  units of heparanase activity, while in the presence of the C-terminal peptide antisera, only  $0.8 \pm 0.2$  units of heparanase were detected ( $p < 0.001$ ; results confirmed in a second experiment). The N-terminal peptide antiserum was not able to neutralize the heparanase activity. Similar results were obtained when the ability of the synthetic peptides to neutralize heparanase activity was examined. Heparanase assays conducted with 3 nM enzyme, 47 nM  $^{35}$ S-HSPG substrate, and varying concentrations of peptides showed that heparanase activity was only 5% of control values in the presence of 250  $\mu$ M C-terminal peptide. By contrast, heparanase activity in the presence of 250  $\mu$ M of either the N-terminal peptide or an unrelated peptide (PLALWAR) was 67% of control values. The ability of both the C-terminal peptide (SEQ ID NO: 42) or antisera from a chicken immunized with the CTAP-III and NAP-2 possess heparanase activity, and suggests that the C-terminal region is essential for catalysis. Modeling of this domain (SEQ ID NO: 42) can be used in the identification of potent peptide-mimetic compounds capable of inhibiting this enzyme activity.

- Computer assisted modeling can be accomplished using programs for automated docking of molecules within 3D databases, as described in DesJarlais, R.L., Sheridan, R.P., Seibel, G.L., Dixon, J.S., Kuntz, I.D., Venkataraghavan, R., "Using shape complementarity as an initial screen in designing ligands for a receptor binding site of known three-dimensional structure"; J. Med. Chem. 31:722-729, 1988. Also, automated de novo construction of ligands that can bind the catalytic site as described in Moon, J.B., Howe, W.J., "Computer design of bioactive molecules: a method for receptor-based de novo ligand design"; Proteins: Struct., Funct., and Genetics, 11:314-328, 1981.

-33-

## SEQUENCE LISTING

## 5 (1) GENERAL INFORMATION:

- (i) APPLICANT: Hoogewerf, Arlene J.  
Ledbetter, Steven R.
- 10 (ii) TITLE OF INVENTION: USE OF HEPARANASE TO IDENTIFY AND  
ISOLATE ANTI-HEPARANASE COMPOUNDS
- (iii) NUMBER OF SEQUENCES: 43
- 15 (iv) CORRESPONDENCE ADDRESS:  
(A) ADDRESSEE: The Upjohn Company, Intellectual Property Law  
(B) STREET: 301 Henrietta  
(C) CITY: Kalamazoo  
(D) STATE: MI  
20 (E) COUNTRY: USA  
(F) ZIP: 49001
- (v) COMPUTER READABLE FORM:  
(A) MEDIUM TYPE: Floppy disk  
25 (B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25
- (vi) CURRENT APPLICATION DATA:  
30 (A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:
- (vii) ATTORNEY/AGENT INFORMATION:  
35 (A) NAME: Jameson, William G.  
(B) REGISTRATION NUMBER: 27,199  
(C) REFERENCE/DOCKET NUMBER: 4731.1 CP
- (ix) TELECOMMUNICATION INFORMATION:  
40 (A) TELEPHONE: 616/385-7561  
(B) TELEFAX: 616/385-6897  
(C) TELEX: 224401

## 45 (2) INFORMATION FOR SEQ ID NO:1:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 85 amino acids  
50 (B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

60 Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala  
1 5 10 15  
Glu Leu Arg Cys Met Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys  
20 25 30  
65 Asn Ile Gln Ser Leu Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln  
35 40 45  
Val Glu Val Ile Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp  
50 55 60

-34-

Pro Asp Ala Pro Arg Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly  
65 70 75 80

Asp Glu Ser Ala Asp  
85

5

## (2) INFORMATION FOR SEQ ID NO:2:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 255 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

AACTTGGCGA AAGGCAAAGA GGAAAGTCTA GACAGTGA CTGATGCTGA ACTCCGCTGC 60  
ATGTGTATAA AGACAACCTC TGGAATTCAT CCCAAAAACA TCCAAAGTTT GGAAGTGATC 120  
GGGAAAGGAA CCCATTGCAA CCAAGTCGAA GTGATAGCCA CACTGAAGGA TGGGAGGAAA 180  
ATCTGCCTGG ACCCAGATGC TCCAGAATC AAGAAAATTG TACAGAAAAA ATTGGCAGGT 240  
GATGAATCTG CTGAT 255

25

## (2) INFORMATION FOR SEQ ID NO:3:

30

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 94 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

35

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ser Ser Thr Lys Gly Gln Thr Lys Arg Asn Leu Ala Lys Gly Lys Glu  
1 5 10  
Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu Leu Arg Cys Met Cys Ile  
20 25 30  
Lys Thr Thr Ser Gly Ile His Pro Lys Asn Ile Gln Ser Leu Glu Val  
35 40 45  
Ile Gly Lys Gly Thr His Cys Asn Gln Val Glu Val Ile Ala Thr Leu  
50 55 60  
Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro Arg Ile Lys  
65 70 75 80  
Lys Ile Val Gln Lys Lys Leu Ala Gly Asp Glu Ser Ala Asp  
85 90

45

50

55

## (2) INFORMATION FOR SEQ ID NO:4:

- (1) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 282 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

60

65

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

-35-

TCCTCCACCA AAGGACAAAC TAAGAGAAAC TTGGCGAAAG GCAAAGAGGA AAGTCTAGAC 60  
 AGTGACTTGT ATGCTGAACT CCGCTGCATG TGTATAAAGA CAACCTCTGG AATTCATCCC 120  
 5 AAAAACATCC AAAGTTTGA AGTGATCGGG AAAGGAACCC ATTGCAACCA AGTCGAAGTG 180  
 ATAGCCACAC TGAAGGATGG GAGGAAAATC TGCCTGGACC CAGATGCTCC CAGAATCAAG 240  
 AAAATTGTAC AGAAAAAATT GGCAGGTGAT GAATCTGCTG AT 282

10 (2) INFORMATION FOR SEQ ID NO:5:

(1) SEQUENCE CHARACTERISTICS:  
 15 (A) LENGTH: 81 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala Glu Leu Arg Cys  
 1 5 10 15  
 25 Met Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys Asn Ile Gln Ser  
 20 25 30  
 30 Leu Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln Val Glu Val Ile  
 35 40 45  
 Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp Pro Asp Ala Pro  
 50 55 60  
 35 Arg Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly Asp Glu Ser Ala  
 65 70 75 80  
 Asp

40

(2) INFORMATION FOR SEQ ID NO:6:

(1) SEQUENCE CHARACTERISTICS:  
 45 (A) LENGTH: 243 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

50

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

GGCAAAGAGG AAAGTCTAGA CAGTGACTTG TATGCTGAAC TCCGCTGCAT GTGTATAAAG 60  
 55 ACAACCTCTG GAATTCATCC CAAAAACATC CAAAGTTTGG AAGTGATCGG GAAAGGAACC 120  
 CATTGCAACC AAGTCGAAGT GATAGCCACA CTGAAGGATG GGAGGAAAT CTGCCTGGAC 180  
 CCAGATGCTC CCAGAATCAA GAAAATTGTA CAGAAAAAAT TGGCAGGTGA TGAATCTGCT 240  
 60 GAT 243

(2) INFORMATION FOR SEQ ID NO:7:

65 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 69 amino acids  
 (B) TYPE: amino acid  
 (D) TOPOLOGY: linear

-36-

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:7:

5      Glu Leu Arg Cys Met Cys Ile Lys Thr Thr Ser Gly Ile His Pro Lys  
       1                    5                    10                    15  
       Asn Ile Gln Ser Leu Glu Val Ile Gly Lys Gly Thr His Cys Asn Gln  
                           20                    25                    30  
 10      Val Glu Val Ile Ala Thr Leu Lys Asp Gly Arg Lys Ile Cys Leu Asp  
                           35                    40                    45  
       Pro Asp Ala Pro Arg Ile Lys Lys Ile Val Gln Lys Lys Leu Ala Gly  
                           50                    55                    60  
 15      Asp Glu Ser Ala Asp  
                           65

(2) INFORMATION FOR SEQ ID NO:8:

20      (1) SEQUENCE CHARACTERISTICS:  
             (A) LENGTH: 207 base pairs  
             (B) TYPE: nucleic acid  
             (C) STRANDEDNESS: single  
 25            (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:8:

30      GAAGTCCGCT GCATGTGTAT AAAGACAACC TCTGGAATTC ATCCCAAAAA CATCCAAAGT      60  
       TTGGAAGTGA TCGGGAAAGG AACCCATTGC AACCAAGTCG AAGTGATAGC CACACTGAAG      120  
 35      GATGGGAGGA AAATCTGCCT GGACCCAGAT GCTCCCAGAA TCAAGAAAAT TGTACAGAAA      180  
       AAATTGGCAG GTGATGAATC TGCTGAT      207

(2) INFORMATION FOR SEQ ID NO:9:

40      (1) SEQUENCE CHARACTERISTICS:  
             (A) LENGTH: 19 amino acids  
             (B) TYPE: amino acid  
             (D) TOPOLOGY: linear  
 45

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:9:

50      Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala  
       1                    5                    10                    15  
       Glu Leu Arg

55

(2) INFORMATION FOR SEQ ID NO:10:

60      (1) SEQUENCE CHARACTERISTICS:  
             (A) LENGTH: 16 amino acids  
             (B) TYPE: amino acid  
             (D) TOPOLOGY: linear

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:10:

      Ser Ser Thr Lys Gly Gln Thr Lys Arg Asn Leu Ala Lys Gly Lys Glu  
       1                    5                    10                    15

-37-

## (2) INFORMATION FOR SEQ ID NO:11:

## (1) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 17 amino acids  
(B) TYPE: amino acid  
(D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:11:

10 Asn Leu Ala Lys Gly Lys Glu Glu Ser Leu Asp Ser Asp Leu Tyr Ala  
1 5 10 15  
15 Glu

## (2) INFORMATION FOR SEQ ID NO:12:

## (1) SEQUENCE CHARACTERISTICS:

- 20 (A) LENGTH: 101 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
25 (D) TOPOLOGY: linear

## (11) MOLECULE TYPE: peptide

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:12:

30 Met Ser Ser Ala Ala Gly Phe Cys Ala Ser Arg Pro Gly Leu Leu Phe  
1 5 10 15  
35 Leu Gly Leu Leu Leu Leu Pro Leu Val Val Ala Phe Ala Ser Ala Glu  
20 25 30  
Ala Glu Glu Asp Gly Asp Leu Gln Cys Leu Cys Val Lys Thr Thr Ser  
35 40 45  
40 Gln Val Arg Pro Arg His Ile Thr Ser Leu Glu Val Ile Lys Ala Gly  
50 55 60  
45 Pro His Cys Pro Thr Ala Gln Leu Ile Ala Thr Leu Lys Asn Gly Arg  
65 70 75 80  
Lys Ile Cys Leu Asp Leu Gln Ala Pro Leu Tyr Lys Lys Ile Ile Lys  
85 90 95  
50 Lys Leu Leu Glu Ser  
100

## (2) INFORMATION FOR SEQ ID NO:13:

## (1) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 439 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
60 (D) TOPOLOGY: linear

## (x1) SEQUENCE DESCRIPTION: SEQ ID NO:13:

65 CCGCAGCATG AGCTCCGCAG CCGGGTTCTG CGCTCACGC CCCGGGCTGC TGTTCCTGGG 60  
GTTGCTGCTC CTGCCACTTG TGGTCGCCTT CGCCAGCGCT GAAGCTGAAG AAGATGGGGA 120  
CCTGCAGTGC CTGTGTGTGA AGACCACCTC CCAGGTCCGT CCCAGGCACA TCACCAGCCT 180

-38-

GGAGGTGATC AAGGCCGGAC CCCACTGCCC CACTGCCCCA CTGATAGCCA CGCTGAAGAA 240  
 TGGAAGGAAA ATTTGCTTGG ACCTGCAAGC CCCGCTGTAC AAGAAAATAA TTAAGAAACT 300  
 5 TTTGGAGAGT TAGCTACTAG CTGCCTACGT GTGTGCATTT GCTATATAGC ATACTTCTTT 360  
 TTTCCAGTTT CAATCTAACT GTGAAAGAAA CTTCTGATAT TTGTGTTATC CTTATGATTT 420  
 TAAATAAACA AAATAAATC 439

10

## (2) INFORMATION FOR SEQ ID NO:14:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 101 amino acids  
 15 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide  
 20

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

25 Met Asn Gln Thr Ala Ile Leu Ile Cys Cys Leu Ile Phe Leu Thr Leu  
 1 5 10 15  
 Ser Gly Ile Gln Gly Val Pro Leu Ser Arg Thr Val Arg Cys Thr Cys  
 20 25 30  
 30 Ile Ser Ile Ser Asn Gln Pro Val Asn Pro Val Asn Pro Arg Ser Leu  
 35 40 45  
 Glu Lys Leu Glu Ile Ile Pro Ala Ser Gln Phe Cys Pro Arg Val Glu  
 50 55 60  
 35 Ile Ile Ala Thr Met Lys Lys Lys Gly Glu Lys Arg Cys Leu Asn Pro  
 65 70 75 80  
 40 Glu Ser Lys Ala Ile Lys Asn Leu Leu Lys Ala Val Ser Lys Glu Met  
 85 90 95  
 Ser Lys Arg Ser Pro  
 100

45

## (2) INFORMATION FOR SEQ ID NO:15:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 650 base pairs  
 50 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

55

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

GAGACATTCC TCAATTGCTT AGACATATTC TGAGCCTACA GCAGAGGAAC CTCCAGTCTC 60  
 60 AGCACCATGA ATCAAAGTGC GATTCTGATT TGCTGCCTTA TCTTTCTGAC TCTAAGTGGC 120  
 ATTCAAGGAG TACCTCTCTC TAGAACCGTA CGCTGTACCT GCATCAGCAT TAGTAATCAA 180  
 CCTGTTAATC CAAGGTCTTT AGAAAACTT GAAATTATTC CTGCAAGCCA ATTTTGTCCA 240  
 65 CGTGTGAGAG TCATTGCTAC AATGAAAAAG AAGGGTGAGA AGAGATGTCT GAATCCAGAA 300  
 TCGAAGGCCA TCAAGAATTT ACTGAAAGCA GTTAGCAAGG AAATGTCTAA AAGATCTCCT 360



-39-

TAAAACCAGA GGGGAGCAAA ATCGATGCAG TGCTTCCAAG GATGGACCAC ACAGAGGCTG 420  
 CCTCTCCCAT CACTTCCCTA CATGGAGTAT ATGTCAAGCC ATAATTGTTC TTAGTTTGCA 480  
 5 GTTACTACTAA AAGGTGACCA ATGATGGTCA CCAAATCAGC TGCTACTACT CCTGTAGGAA 540  
 GGTAAATGTT CATCATCCTA AGCTATTAGC TAATAACTCT ACCCTGGCAC TATAATGTAA 600  
 GCTCTACTGA GGTGCTATGT TCTTAGTGGA TGTTCTGACC CTGCTTCAAA 650

## (2) INFORMATION FOR SEQ ID NO:16:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 107 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Met Ala Arg Ala Ala Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu  
 1 5 10 15  
 Arg Val Ala Leu Leu Leu Leu Leu Val Ala Ala Gly Arg Arg Ala  
 20 25 30  
 Ala Gly Ala Ser Val Ala Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr  
 35 40 45  
 Leu Gln Gly Ile His Pro Lys Asn Ile Gln Ser Val Asn Val Lys Ser  
 50 55 60  
 Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn  
 65 70 75 80  
 Gly Arg Lys Ala Cys Leu Asn Pro Ala Ser Pro Ile Val Lys Lys Ile  
 85 90 95  
 Ile Glu Lys Met Leu Asn Ser Asp Lys Ser Asn  
 100 105

## (2) INFORMATION FOR SEQ ID NO:17:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1050 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CTCGCCAGCT CTTCCGCTCC TCTCACAGCC GCCAGACCCG CCTGCTGAGC CCCATGGCCC 60  
 60 GCGTGCTCT CTCCGCCGCC CCCAGCAATC CCCGGCTCCT GCGAGTGGCA CTGCTGCTCC 120  
 TGCTCCTGGT AGCCGCTGGC CGGCGCGCAG CAGGAGCGTC CGTGGCCACT GAACTGCGCT 180  
 GCCAGTGCTT GCAGACCCTG CAGGGAATTC ACCCCAAGAA CATCAAAGT GTGAACGTGA 240  
 65 AGTCCCCCGG ACCCCACTGC GCCCAAACCG AAGTCATAGC CAACTCAAG AATGGGCGGA 300  
 AAGCTTGCCCT CAATCCTGCA TCCCCCATAG TTAAGAAAAT CATCGAAAAG ATGCTGAACA 360

-40-

GTGACAAATC CAACTGACCA GAAGGGAGGA GGAAGCTCAC TGGTGGCTGT TCCTGAAGGA 420  
 GGCCCTGCCC TTATAGGAAC AGAAGAGGAA AGAGAGACAC AGCTGCAGAG GCCACCTGGA 480  
 5 TTGTGCCTAA TGTGTTTGAG CATCGCTTAG GAGAAGTCTT CTATTATTTT ATTTATTTCAT 540  
 TACTTTTGAA GATTCTATGT TAATATTTTA GGTGTA AAAAT AATTAAGGGT ATGATTA ACT 600  
 CTACCTGCAC ACTGTCCTAT TATATTCATT CTTTTTGAAA TGTCAACCCC AAGTTAGTTC 660  
 10 AATCTGGATT CATATTTAAT TTGAAGGTAG AATGTTTTCA AATGTTCTCC AGTCATTATG 720  
 TTAATATTTT TGAGGAGCCT GCAACATGCC AGCCACTGTG ATAGAGGCTG GCGGATCCAA 780  
 15 GCAAATGGCC AATGAGATCA TTGTGAAGGC AGGGGAATGT ATGTGCACAT CTGTTTTGTA 840  
 ACTGTTTAGA TGAATGTCAG TTGTTATTTA TTGAAATGAT TTCACAGTGT GTGGTCAACA 900  
 TTTCTCATGT TGAACTTTA AGAACTAAAA TGTCTAAAT ATCCCTTGGA CATTTTATGT 960  
 20 CTTTCTTGTA AGGCATACTG CCTTGTTTAA TGGTAGTTTT ACAGTGTTC TGGCTTAGAA 1020  
 CAAAGGGGCT TAATTATTGA TGTTCGGA 1050

## 25 (2) INFORMATION FOR SEQ ID NO:18:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 102 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: peptide

## 35 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

Met Ala Arg Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu  
 1 5 10 15  
 Arg Val Ala Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala  
 20 25 30  
 Ala Gly Ala Pro Lys Ala Thr Glu Lys Arg Cys Gln Cys Lys Gln Thr  
 35 40 45  
 Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Lys Val Lys Ser  
 50 55 60  
 Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn  
 65 70 75 80  
 Gly Gln Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Lys Lys Ile  
 85 90 95  
 Ile Glu Lys Met Leu Lys  
 100

## 60 (2) INFORMATION FOR SEQ ID NO:19:

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1081 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

-41-

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CTCTCCTCCT CGCACAGCCG CTCGAACCGC CTGCTGAGCC CCATGGCCCC CGCCACGCTC 60  
 5 TCCGCCGCCC CCAGCAATCC CCGGCTCCTG CGGGTGGCGC TGCTGCTCCT GCTCCTGGTG 120  
 GCCGCCAGCC GGC GCGCAGC AGGAGCGCCC CTGGCCACTG AACTGCGCTG CCAGTGCTTG 180  
 CAGACCCTGC AGGGAATTCA CCTCAAGAAC ATCCAAAGTG TGAAGGTGAA GTCCCCCGGA 240  
 10 CCCCAC TGCG CCCAAACCGA AGTCATAGCC AACTCAAGA ATGGGCAGAA AGCTTGCTCTC 300  
 AACCCCGCAT CGCCCATGGT TAAGAAAATC ATCGAAAAGA TGCTGAAAAA TGGCAAATCC 360  
 15 AACTGACCAG AAGGAAGGAG GAAGCTTATT GGTGGCTGTT CCTGAAGGAG GCCCTGCCCT 420  
 TACAGGAACA GAAGAGGAAA GAGAGACACA GCTGCAGAGG CCACCTGGAT TGCGCCTAAT 480  
 GTGTTTGAGC ATCACTTAGG AGAAGTCTTC TATTTATTTA TTTATTTATT TATTTGTTTG 540  
 20 TTTTAGAAGA TTCTATGTTA ATATTTTATG TGTAATAATA GGTTATGATT GAATCTACTT 600  
 GCACACTCTC CCATTATATT TATTGTTTAT TTTAGGTCAA ACCCAAGTTA GTTCAATCCT 660  
 25 GATTCATATT TAATTTGAAG ATAGAAGGTT TGCAGATATT CTCTAGTCAT TTGTTAATAT 720  
 TTCTTCGTGA TGACATATCA CATGTCAGCC ACTGTGATAG AGGCTGAGGA ATCCAAGAAA 780  
 ATGGCCAGTG AGATCAATGT GACGGCAGGG AAATGTATGT GTGTCTATTT TGTAAGTGTA 840  
 30 AAGATGAATG TCAGTTGTTA TTTATTGAAA TGATTTTACA GTGTGTGGTC AACATTTCTC 900  
 ATGTTGAAGC TTTAAGAACT AAAATGTTCT AAATATCCCT TGGACATTTT ATGTCTTTCT 960  
 35 TGTAAGGCAT ACTGCCTTGT TTAATGTTAA TTATGCAGTG TTTCCCTCTG TGTTAGAGCA 1020  
 GAGAGGTTTC GATATTTATT GATGTTTTCA CAAAGAACAG GAAAATAAAA TATTTAAAAA 1080  
 T 1081

(2) INFORMATION FOR SEQ ID NO:20:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 107 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(11) MOLECULE TYPE: peptide

(x1) SEQUENCE DESCRIPTION: SEQ ID NO:20:

55 Met Ala His Ala Thr Leu Ser Ala Ala Pro Ser Asn Pro Arg Leu Leu  
 1 5 10 15  
 Arg Val Ala Leu Leu Leu Leu Leu Val Ala Ala Ser Arg Arg Ala  
 20 25 30  
 60 Ala Gly Ala Ser Val Val Thr Glu Leu Arg Cys Gln Cys Leu Gln Thr  
 35 40 45  
 Leu Gln Gly Ile His Leu Lys Asn Ile Gln Ser Val Asn Val Arg Ser  
 50 55 60  
 Pro Gly Pro His Cys Ala Gln Thr Glu Val Ile Ala Thr Leu Lys Asn  
 65 70 75 80

-42-

Gly Lys Lys Ala Cys Leu Asn Pro Ala Ser Pro Met Val Gln Lys Ile  
                                     85                                    90                                    95

Ile Glu Lys Ile Leu Asn Lys Gly Ser Thr Asn  
                                     100                                    105

5

## (2) INFORMATION FOR SEQ ID NO:21:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 988 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

10

15

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

CTCGCACAGC TTCCCCGACGC GTCTGCTGAG CCCCATGGCC CACGCCACGC TCTCCGCCGC 60  
 20 CCCCAGCAAT CCCC GGCTCC TCGGGGTGGC GCTGCTGCTC CTGCTCCTGG TGGCCGCCAG 120  
 CCGGCGCGCA GCAGGAGCGT CCGTGGTCAC TGAAGTGGC TGCCAGTGCT TGCAGACACT 180  
 25 GCAGGGAATT CACCTCAAGA ACATCCAAAG TGTGAATGTA AGGTCCCCCG GACCCCACTG 240  
 CGCCCCAACC GAAGTCATAG CCACACTCAA GAATGGGAAG AAAGCTTGTC TCAACCCCGC 300  
 ATCCCCCATG GTTCAGAAAA TCATCGAAAA GATACTGAAC AAGGGGAGCA CCAACTGACA 360  
 30 GGAGAGAAGT AAGAAGCTTA TCAGCGTATC ATTGACACTT CCTGCAGGGT GGTCCCTGCC 420  
 CTTACCAGAG CTGAAAATGA AAAAGAGAAC AGCAGCTTTC TAGGGACAGO TGGAAAGGAG 480  
 35 TTAATGTGTT TGAATATTTT TTACGAGGGT TCTACTTATT TATGTATTTA TTTTGTAAAG 540  
 CTTGTATTTT AATATTTTAC ATGCTGTTAT TTAAAGATGT GAGTGTGTTT CATCAAACAT 600  
 AGCTCAGTCC TGATTATTTA ATTGGAATAT GATGGGTTTT AAATGTGTCA TTAAACTAAT 660  
 40 ATTTAGTGGG AGACCATAAT GTGTCAGCCA CCTTGATAAA TGACAGGGTG GGGAACTGGA 720  
 GGGTGGGGGG ATTGAAATGC AAGCAATTAG TGGATCACTG TTAGGGTAAG GGAATGTATG 780  
 45 TACACATCTA TTTTATATAC TTTTATTTTA AAAAAAGAAT GTCAGTTGTT ATTTATTCAA 840  
 ATTATCTCAC ATTATGTGTT CAACATTTT ATGCTGAAGT TTCCCTTAGA CATTTTATGT 900  
 CTTGCTTGTA GGCATAATG CCTTGTTTAA TGTCCATTCT GCAGCGTTTC TCTTCCCTT 960  
 50 GGAAAAGAGA ATTTATCATT ACTGTTAC 988

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 97 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

55

60

## (ii) MOLECULE TYPE: peptide

65

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Met Thr Ser Lys Leu Ala Val Ala Leu Leu Ala Ala Phe Leu Ile Ser  
                                     5                                    10                                    15

-43-

Ala Ala Leu Cys Glu Gly Ala Val Leu Pro Arg Ser Ala Lys Glu Leu  
                   20                  25                  30  
 5 Arg Cys Gln Cys Ile Lys Thr Tyr Ser Lys Pro Phe His Pro Lys Phe  
                   35                  40                  45  
 Ile Lys Glu Leu Arg Val Ile Glu Ser Gly Pro His Cys Ala Asn Thr  
                   50                  55                  60  
 10 Glu Ile Ile Val Lys Leu Ser Asp Gly Arg Glu Leu Cys Leu Asp Pro  
                   65                  70                  75                  80  
 Lys Glu Asn Trp Val Gln Arg Val Val Glu Lys Phe Leu Lys Arg Ala  
                   85                  90                  95  
 15 Glu

## (2) INFORMATION FOR SEQ ID NO:23:

- 20 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 291 base pairs  
 (B) TYPE: nucleic acid  
 25 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

30 ATGACTTCCA AGCTGGCCGT GGCTCTCTTG GCAGCCTTCC TGATTCTGCG AGCTCTGTGT 60  
 GAAGGTGCAG TTTTGCCAAG GAGTGCTAAA GAACTTAGAT GTCAGTGCAT AAAGACATAC 120  
 35 TCCAAACCTT TCCACCCCAA ATTTATCAAA GAACTGAGAG TGATTGAGAG TGGACCACAC 180  
 TCGGCCAACA CAGAAATTAT TGTAAGCTT TCTGATGGAA GAGAGCTCTG TCTGGACCCC 240  
 40 AAGGAAACT GGGTGCAGAG GGTGTGGAG AAGTTTTTGA AGAGGGCTGA G 291

## (2) INFORMATION FOR SEQ ID NO:24:

- 45 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 78 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

## (11) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

55 Ala Gly Pro Ala Ala Ala Val Leu Arg Glu Lys Arg Cys Val Cys Leu  
                   1                  5                  10                  15  
 Gln Thr Thr Gln Gly Val His Pro Lys Met Ile Ser Asn Leu Gln Val  
                   20                  25                  30  
 60 Phe Ala Ile Gly Pro Gln Cys Ser Lys Val Glu Val Val Ala Ser Leu  
                   35                  40                  45  
 Lys Asn Gly Lys Glu Ile Cys Leu Asp Pro Glu Ala Pro Phe Leu Lys  
                   50                  55                  60  
 65 Lys Val Ile Gln Lys Ile Leu Asp Gly Gly Asn Lys Glu Asn  
                   65                  70                  75

-44-

## (2) INFORMATION FOR SEQ ID NO:25:

## (1) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 216 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GTGTTGCGGG AACTGCGGTG CGTGTGTTTA CAGACCACGC AGGGAGTTCA TCCCAAAATG 60  
15 ATCAGTAATC TGCAAGTGTT CGCCATAGGC CCACAGTGCT CCAAGGTGGA AGTGGTAGCC 120  
TCCCTGAAGA ACGGGAAGGA AATTTGTCTT GATCCAGAAG CCCCTTTTCT AAAGAAAGTC 180  
ATCCAGAAAA TCCTCGACGG CGGCAACAAA GAAAAC 216

20

## (2) INFORMATION FOR SEQ ID NO:26:

## (i) SEQUENCE CHARACTERISTICS:

- 25 (A) LENGTH: 93 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

30

## (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

35 Met Gln Val Ser Thr Ala Ala Leu Ala Val Leu Leu Cys Thr Met Ala  
1 5 10 15  
Leu Cys Asn Gln Val Leu Ser Ala Pro Leu Ala Ala Asp Thr Pro Thr  
20 25 30  
40 Ala Cys Cys Phe Ser Tyr Thr Ser Arg Gln Ile Pro Gln Asn Phe Ile  
35 40 45  
Ala Asp Tyr Phe Glu Thr Ser Ser Gln Cys Ser Lys Pro Ser Val Ile  
50 55 60  
Phe Leu Thr Lys Arg Gly Arg Gln Val Cys Ala Asp Pro Ser Glu Glu  
65 70 75 80  
50 Trp Val Gln Lys Tyr Val Ser Asp Leu Glu Leu Ser Ala  
85 90

## (2) INFORMATION FOR SEQ ID NO:27:

## (1) SEQUENCE CHARACTERISTICS:

- 55 (A) LENGTH: 4788 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
60 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

65 GAATTCAAGG CCTGTCTGG TTTGGTCCCA ATTTACCTTT ATCATCCATA TTCACCCCCA 60  
CTGCTCTGCA GCTCCACTGA AGCACCCCTT CTTTCTCTG AGCCACAATG TCACACCCAG 120  
GACTCTGCCT CAGCTGGGCC TCCACTGCCC ACCCATCTAT AGATGCCTAA ATCCCGGGCA 180

	GTTATCCAGA	CACAACTAAA	GTTCCATCCC	TTCCATGAAG	CCTTCCCCAA	CCCTCTGGTG	240
	GAAGGTCACT	TCTTCCTCAT	GGGGTTCTGA	GCTTTCATTT	CTTTTCTAC	TAAGAGTTTT	300
5	ACAATTACCT	GTTCATACAC	TCTACCTGCC	CCCATGAGAC	CAGGGGCATC	TCAGAAACAA	360
	AGATCATTAA	AACCAACTAA	ATCTATTTCT	CATTATAAAA	TGAGATATGC	TGATTGATTG	420
	CAAAATAATA	AAATAACAAA	GTATGGAAAA	GAAAAAATAA	AGCATATAAT	CTGGCTGAGA	480
10	AGGTAGAGAC	CCTTCCACAC	CACTGAAATT	ATGTGTTGAA	AAGAATAAGG	AAAAAACTGC	540
	TTCAGTTTGG	CATTATTTAT	GTAAGTATAG	TATAGGATCC	TTAAATGGT	TCAAAGAAAT	600
15	GGGAAATCAA	GACTTCATTT	TGGCAAAGCC	ATTGAACAGA	AACTGTAGCA	TATTTATCAG	660
	TAATTTCTTT	CAGATTAAAC	AACTGACAAC	AACCCACTTT	TCAACCAGTG	ATGTTGGAAA	720
	TGTTTTAAAA	CAAAATTAGT	TCATAAATTT	GTGGGTTGAC	CAAGAAGGTA	ATAAAGTCTC	780
20	ACTAAATAAA	ATGAGGAAAA	TTCAGAAAAA	GAAAAAATA	AGAAAATAAA	TCACCCATGG	840
	ATCTAAGCAC	TATTCATTCT	TTAAGGCATG	TATTTCCAAG	CCTTTTAATT	TTTTCATGCC	900
25	TAGAGTTGGC	ATGGCATATA	TATATCTTTA	TACAATTCTT	CAAATTTTAT	AGAATTTGTA	960
	TAATGTTTTA	TCTTGCTTTT	TTTTTAACCA	CTGATGTTAT	AAGCATATTT	ATGCCACTTC	1020
	ATTCACGTTA	GAGACTTAAT	AATAAAGGAT	CTTGTTGATA	ATTTATCATT	CCCTGATAGA	1080
30	GAAAAATTTA	GCTTTGCTTA	TTTTAGAGTT	ATAAATGATG	CTGGGTCAGG	TATCTTTATG	1140
	TTTGAAGATG	GCTCCATATT	TGGGTTGTTT	CCACAGAACT	CTTCCAGAA	ATGCTTTTTT	1200
35	TAGGTTAATG	GCTACACATA	TTTCTAGGCA	CCTGACATAC	TGACACCCAC	CTCTAAAGTA	1260
	TTTTTATGAT	CCACAAC TAG	CGTTTAACAC	AGCGCCCCAG	TCACTCCGAG	ACTAATAAAT	1320
	AGACAAATGA	CTGAAACGTG	ACCTCATGCT	TTCTATTCC	CCAGCTTTCA	TTGAGTTCCT	1380
40	TTCTCTGGG	AGGACTGGGG	GTTGTCTAGC	CCTCCACAGC	ATCAGCCCAT	TGACCCATATC	1440
	CTTGTTGGTTA	TAGCAGCTGA	GGAAGCAGAA	TTACAGCTCT	GTGGGAAGGA	ATGGGGCTGG	1500
45	AGAGTTCATG	CATAGACCAA	TTCTTTTTTT	TTTTTTTTTT	TGAGATGGAG	TTTCACTTTT	1560
	GTTGCCCAGG	CTGGAGTGCA	ATGGCATGAT	CTCAGCTCAC	CACAGCCCCC	ACCTCCTGGG	1620
	TTCAAGCGAT	TCTCCTGCCC	TCAGCCTCCC	GAGTAGCTGG	GATTACAGGC	ATGTGCCACC	1680
50	ACGCCTGACT	ACTTTTGTAT	TTTTAGTAGA	GATGGAGTTT	CTCTTCTTG	GTCAGGTTGG	1740
	TCTCAAACCTC	CTGACCTCAG	GTGATCTGCA	GCCTCGGCCT	CCAAAGTGTT	GGGATTACAG	1800
55	GTGTGAGCGA	CCATGCCTGG	CTGCATAGAC	CAGTTCTTAT	GAGAAGGGAT	CAACTAAGAA	1860
	TAGCCTTGGG	TTGACACACA	CCCCTCTTCA	CACTCACAGG	AGAAACCCCA	TGAAGCTAGA	1920
	ACCAGTCATG	AGTTGAGAGC	TGAGAGTTAG	AGAGTAGCTC	AGAGATGCTA	TTCTTGGATA	1980
60	TCCTGAGCCC	CTGTGGTCAC	CAGGGACCC	GAGTTGTGCA	AACTCAGCA	TGACAGCATC	2040
	ACTACACTTA	AAAATTTCCC	TCCTCACCCC	CAGATTCCAT	TTCCCCATCC	GCCAGGGCTG	2100
65	CCTATAAAGA	GGAGAGATGG	CTTCAGACAT	CAGAAGGACG	CAGGCAGCAA	AGAGTAGTCA	2160
	GTCCCTTCTT	GGCTCTGCTG	ACACTCGAGC	CCACATTCCA	TCACCTGCTC	CCAATCATGC	2220
	AGGTCTCCAC	TGCTGCCCTT	GCCGTCCTCC	TCTGCACCAT	GGCTCTCTGC	AACCAGSTCC	2280

	TCTCTGCACC	ACGTGAGTCC	ATGTTGTTGT	TGTGGGTATC	ACCACTCTCT	GGCCATGGTT	2340
	AGACCACATC	AGTCTTTTTT	TGTGGCGTGA	GAGGCCCCGA	AGAGAAAAGA	AGGAAGTTCT	2400
5	TAAAGCGCTG	CCAAACACCT	TGGTCTTTTT	CTTCACAAC	TTTATTTT	TCTCTAGAAG	2460
	GGGTCTTAGC	CCTCCTAGTC	TCCAGGTATG	AGAATCTAGG	CAGGGGCAGG	GGAGTTACAG	2520
	TCCCTTGTA	AGATAGAAAA	ACAGGGTTCA	AAACGAATCA	GTTTGCAAGA	GGCAGAATCC	2580
10	AGGGCTGCTT	ACTTCCCAGT	GGGGTCTGTT	CTTCACTCTC	CAGCTCACCC	TAGTCTCCCA	2640
	GGAGCCCTGT	CCCTTGGATG	TCTTATGAGA	GATGTCCAGG	GCTTCTCTTG	GGCTGGGGTA	2700
15	TGACTTCTTG	AACCGACAAA	ATTCCATGAA	GAGAGCTAAG	AGAACAGTCC	ATTCAGGTAT	2760
	CTGGATCACA	TAGAGAAACA	GAGAACCCAC	TATGAAGAGT	CAAGGGGAAA	GAGGAATATA	2820
	GACAGAAACA	AAGAGACATT	TCTCTGCAAA	ACCCCCAAA	TGCCTTGCA	TCACTTGGTC	2880
20	TGAGCAAGCC	TGCCCTCCTC	AACCACTCAG	GGATCAGAAG	CTGCCTGGCC	TTTTCTTCTG	2940
	AGCTGTGACT	TGGGCTTATT	CTCTCCTTTC	TCCGCAGTTG	CTGCTGACAC	GCCGACCGCC	3000
25	TGCTGCTTCA	GCTACACCTC	CCGACAGATT	CCACAGAATT	TCATAGCTGA	CTACTTTGAG	3060
	ACGAGCAGCC	AGTGCTCCAA	GCCCAGTGTC	ATGTAAGTGC	CAGTCTTCCT	GCTCACCTCT	3120
	AGGGAGGTAG	GGAGTGTCAG	GGTGGGGGCA	GAAACAGGCC	AGAAGGCCAT	CCTGGAAGG	3180
30	CCCAGCCTTC	AGGAGCCTAT	CGGGGATACA	GGACGCAGGG	CACTGAGGTG	TGACCTGACT	3240
	TGGGGCTGGA	GTGAGGTGGG	TGTTACAGAG	TCAGGAAGGG	CTGCCCCAGG	CCAGAGGAAA	3300
35	GGGACAGGAA	GAAGGAGGCA	GCAGGACACT	CTGAGGGCCC	CCTTGCCTGG	AGTCACTGAG	3360
	AGAAGCTCTC	TAGACGGAGA	TAGGCAGGGG	GCCCTGAGA	GAGGAGCAGG	CCTTGAGCTG	3420
	CCCAGGACAG	AGAGCAGGAT	GTCAGGGCCA	TGGTGGGCCC	AGGATTCCCC	GGCTGGATTC	3480
40	CCCAGTGCTT	AACCTTCTCT	CCCTTCTCCA	CAGCTTCTTA	ACCAAGAGAG	GCCGGCAGGT	3540
	CTGTGCTGAC	CCCAGTGAGG	AGTGGGTCCA	GAAATACGTC	AGTGACCTGG	AGCTGAGTGC	3600
45	CTGAGGGGTC	CAGAAGCTTC	GAGGCCCAGC	GACCTCAGTG	GGCCCAGTGG	GGAGGAGCAG	3660
	GAGCCTGAGC	CTTGGGAACA	TGCGTGTGAC	CTCCACAGCT	ACCTCTTCTA	TGGACTGGTT	3720
	ATTGCCAAAC	AGCCACACTG	TGGGACTCTT	CTTAACCTAA	ATTTTAATTT	ATTTATACTA	3780
50	TTTAGTTTTT	ATAATTTATT	TTTGATTTCA	CAGTGTGTTT	GTGATTGTTT	GCTCTGAGAG	3840
	TTCCCCCTGT	CCCCTCCACC	TTCCCTCACA	GTGTGTCTGG	TGACAACCGA	GTGGCTGTCA	3900
55	TGGGCTGTG	TAGGCAGTCA	TGGCACCAAA	GCCACCAGAC	TGACAAATGT	GTATCAGATG	3960
	CTTTTGTTCA	GGGCTGTGAT	CGGCCTGGGG	AAATAATAAA	GATGTTCTTT	TAAACGGTAA	4020
	ACCAGTATTG	AGTTTGTTTT	TGTTTTTCTG	GCAATCAAAA	ATCACTGGTT	AAGAGGAATC	4080
60	ATAGGCAAAG	ATTAGGAAGA	GGTGAAATGG	AGGGAAATG	GGAGAGATGG	GGAGCGCTGC	4140
	GACAGAGTTA	TCCACTTCAC	AAAATTCTGG	AACATTGAAA	CTACGAATAT	GTTATAACTC	4200
65	AAATCGTAAT	ATGCACGCTC	TAGGAGAATT	AACTACTTGA	ATGGCCACCA	TTAAGCAGAG	4260
	TATTCTGTAG	GGCATATTCA	TGATGAATCA	AGCTCTTAAT	AGCAATTATT	TACATTGTTG	4320
	AGGCTTACTC	CTCCTACTGA	GTGCTTTTTA	TACATTGTTC	ATTTAATCTT	ACCAATGCAA	4380



-47-

TAGTACAGCT TAGGTACTAT TAATACCTCC ACTTGACAGA AAAGTAACCC AGGGCTCAGA 4440  
 AAGGTTAGAC AACTTGGCTG AGGTACACA GCACGTAAAC GGTCAATTGT GTTCCAAAAC 4500  
 5 TGGACTTTTA TTGAAC TACA GACTATGCTG TTAACCATTG ACCAAGTTAT TTCCCAAAGT 4560  
 ATGACCCGCC TATACTCAA TCTTACCCCA TTCTTTAACA GATGATACTT TATCCATTGC 4620  
 AACCATTCC TGTCAGGATT CTGAGTTGAC ATAGAGTGTT TCAGCAGTGA TTATTTAAGC 4680  
 10 CAATTACATC AGGATCTTTA GGTGTAGACC TGGGAAGTGA TATTTTATC AAGCTCATGA 4740  
 GGTGTTCCAT AGCATGTTAA TGAAGTGAAG CCACTGTCAA TAGAATTC 4788

15 (2) INFORMATION FOR SEQ ID NO:28:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 92 amino acids  
 (B) TYPE: amino acid  
 20 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

25

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

Met Lys Leu Cys Val Thr Val Leu Ser Leu Leu Met Leu Val Ala Ala  
 30 1 5 10 15  
 Phe Cys Ser Pro Ala Leu Ser Ala Pro Met Gly Ser Asp Pro Pro Thr  
 20 25 30  
 35 Ala Cys Cys Phe Ser Tyr Thr Ala Arg Lys Leu Pro Arg Asn Phe Val  
 35 40 45  
 Val Asp Tyr Tyr Glu Thr Ser Ser Leu Cys Ser Gln Pro Ala Val Val  
 40 50 55 60  
 Phe Gln Thr Lys Arg Ser Lys Gln Val Cys Ala Asp Pro Ser Glu Ser  
 65 70 75 80  
 45 Trp Val Gln Glu Tyr Val Tyr Asp Leu Glu Leu Asn  
 85 90

(2) INFORMATION FOR SEQ ID NO:29:

(1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 696 base pairs  
 (B) TYPE: nucleic acid  
 50 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTCCCCCCCC CCCCCCCCCC CCCC GCCCGA GCACAGGACA CAGCTGGGTT CTGAAGCTTC 60  
 60 TGAGTTCTGC AGCCTCACCT CTGAGAAAAC CTCTTTTCCA CCAATACCAT GAAGCTCTGC 120  
 GTGACTGTCC TGTCTCTCCT CATGCTAGTA GCTGCCTTCT GCTCTCCAGC GCTCTCAGCA 180  
 65 CCAATGGGCT CAGACCTCC CACCGCTGC TGCTTTTCTT ACACCGCGAG GAAGCTTCCT 240  
 CGCAACTTTG TGGTAGATTA CTATGAGACC AGCAGCTCT GCTCCAGCC AGCTGTGGTA 300  
 TTCAAACCA AAAGAAGCAA GCAAGTCTGT GCTGATCCCA GTGAATCCTG GGTCCAGGAG 360

-49-

GGGCTCTGGA AACCACATGG CTTCACCTGT CCCCAGAACT ACCAGCCCTA CACCATTCTT 420  
 TCTGCCCTGC TTTTGCTAGG TCACAGAGGA TCTGCTTGGT CTTGATAAGC TATGTTGTTG 480  
 5 CACTTTAAAC ATTTAAATTA TACAATCATC AACCCCAAC 520

## (2) INFORMATION FOR SEQ ID NO:32:

(1) SEQUENCE CHARACTERISTICS:  
 10 (A) LENGTH: 99 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

15 (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

20 Met Lys Val Ser Ala Ala Leu Leu Cys Leu Leu Leu Ile Ala Ala Thr  
 1 5 10 15  
 25 Phe Ile Pro Gln Gly Lys Ala Gln Pro Asp Ala Ile Asn Ala Pro Val  
 20 25 30  
 Thr Cys Cys Tyr Asn Phe Thr Asn Arg Lys Ile Ser Val Gln Arg Leu  
 35 40 45  
 30 Ala Ser Tyr Arg Arg Ile Thr Ser Ser Lys Cys Pro Lys Glu Ala Val  
 50 55 60  
 35 Ile Phe Lys Thr Ile Val Ala Lys Glu Ile Cys Ala Asp Pro Lys Gln  
 65 70 75 80  
 Lys Trp Val Gln Asp Ser Met Asp His Leu Asp Lys Gln Thr Gln Thr  
 85 90 95  
 40 Pro Lys Thr

## (2) INFORMATION FOR SEQ ID NO:33:

45 (1) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 725 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

50

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CTAACCCAGA AACATCCAAT TCTCAAAGT AAGCTCGCAC TCTCGCCTCC AGCATGAAAG 60  
 55 TCTCTGCCGC CCTTCTGTGC CTGCTGCTCA TAGCAGCCAC CTTTATTCCC CAAGGGCTCG 120  
 CTCAGCCAGA TGCAATCAAT GCCCAGTCA CCTGCTGTTA TAACCTCACC AATAGGAAGA 180  
 60 TCTCAGTGCA GAGGCTCGCG AGCTATAGAA GAATCACCAG CAGCAAGTGT CCCAAAGAAG 240  
 CTGTGATCTT CAAGACCATT GTGGCCAAGG AGATCTGTGC TGACCCCAAG CAGAAGTGGG 300  
 TTCAGGATTC CATGGACCAC CTGGACAAGC AAACCCAAAC TCCGAAGACT TGAACACTCA 360  
 65 CTCCACAACC CAAGAATCTG CAGCTAACTT ATTTTCCCCT AGCTTTCCCC AGACACCCTG 420  
 TTTTATTTTA TTATAATGAA TTTTGTGTTG TGATGTGAAA CATTATGCCT TAAGTAATGT 480

-50-

TAATTCCTAT TTAAGTTATT GATGTTTAA GTTATCTTT CATGGTACTA GTGTTTTTTA 540  
 GATACAGAGA CTTGGGGAAA TTGCTTTTCC TCTTGAACCA CAGTTCTACC CCTGGGATGT 600  
 5 TTTGAGGGTC TTGCAAGAA TCATTAATAC AAAGAATTTT TTTAACATT CCAATGCATT 660  
 GCTAAATAT TATTGTGGAA ATGAATATT TGTAATATT ACACCAAATA AATATATTT 720  
 TGTAC 725

10 (2) INFORMATION FOR SEQ ID NO:34:

15 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 99 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

20 (ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

25 Met Lys Ala Ser Ala Ala Leu Leu Cys Leu Leu Leu Thr Ala Ala Ala  
 1 5 10 15  
 Phe Ser Pro Gln Gly Leu Ala Gln Pro Val Gly Ile Asn Thr Ser Thr  
 20 25 30  
 30 Thr Cys Cys Tyr Arg Phe Ile Asn Lys Lys Ile Pro Lys Gln Arg Leu  
 35 40 45  
 Glu Ser Tyr Arg Arg Thr Thr Ser Ser His Cys Pro Arg Glu Ala Val  
 50 55 60  
 35 Ile Phe Lys Thr Lys Leu Asp Lys Glu Ile Cys Ala Asp Pro Thr Gln  
 65 70 75 80  
 40 Lys Trp Val Gln Asp Phe Met Lys His Leu Asp Lys Lys Thr Gln Thr  
 85 90 95  
 Pro Lys Leu

45 (2) INFORMATION FOR SEQ ID NO:35:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 810 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

55 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

AGCAGAGGGG CTGAGACCAA ACCAGAAACC TCCAATTCTC ATGTGGAAGC CCATGCCCTC 60  
 60 ACCCTCCAAC ATGAAAGCCT CTGCAGCACT TCTGTGTCTG CTGCTCACAG CAGCTGCTTT 120  
 CAGCCCCCAG GGGCTTGCTC AGCCAGTTGG GATTAATACT TCAACTACCT GCTGCTACAG 180  
 ATTTATCAAT AAGAAAATCC CTAAGCAGAG GCTGGAGAGC TACAGAAGGA CCACCAGTAG 240  
 65 CCACTGTCCC CGGGAAGCTG TAATCTTCAA GACCAAACCTG GACAAGGAGA TCTGTGCTGA 300  
 CCCCACACAG AAGTGGGTCC AGGACTTTAT GAAGCACCTG GACAAGAAAA CCCAAACTCC 360

-51-

AAAGCTTTGA ACATTCATGA CTGAACTAAA AACAAGCCAT GACTTGAGAA ACAAATAATT 420  
 TGTATACCCT GTCCTTTCTC AGAGTGGTTC TGAGATTATT TTAATCTAAT TCTAAAGAAAT 480  
 5 ATGAGCTTTA TGTAATAATG TGAATCATGG TTTTCTTAG TAGATTTTAA AAGTTATTAA 540  
 TATTTTAATT TAATCTTCCA TGGATTTTGG TGGGTTTGA ACATAAAGCC TTGGATGTAT 600  
 ATGTCATCTC AGTGCTGTAA AACTGTGGG ATGCTCCTCC CTTCTCTACC TCATGGGGGT 660  
 10 ATTGTATAAG TCCTTGCAAG AATCAGTGCA AAGATTTGCT TTAATTGTTA AGATATGATG 720  
 TCCCTATGGA AGCATATTGT TATTATATAA TTACATATTT GCATATGTAT GACTCCCAA 780  
 15 TTTTCACATA AAATAGATTT TTGTAAAAA 810

## (2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:  
 20 (A) LENGTH: 91 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

25 (ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

30 Met Lys Val Ser Ala Ala Arg Leu Ala Val Ile Leu Ile Ala Thr Ala  
 1 5 10 15  
 35 Leu Cys Ala Pro Ala Ser Ala Ser Pro Tyr Ser Ser Asp Thr Thr Pro  
 20 25 30  
 Cys Cys Phe Ala Tyr Ile Ala Arg Pro Leu Pro Arg Ala His Ile Lys  
 35 40 45  
 40 Glu Tyr Phe Tyr Thr Ser Gly Lys Cys Ser Asn Pro Ala Val Val Phe  
 50 55 60  
 Val Thr Arg Lys Asn Arg Gln Val Cys Ala Asn Pro Glu Lys Lys Trp  
 65 70 75 80  
 45 Val Arg Glu Tyr Ile Asn Ser Leu Glu Met Ser  
 85 90

## (2) INFORMATION FOR SEQ ID NO:37:

50 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 1160 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 55 (D) TOPOLOGY: linear

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

60 CCTCCGACAG CCTCTCCACA GGTACCATGA AGGTCTCCGC GGCACGCCTC GCTGTCATCC 60  
 TCATTGCTAC TGCCCTCTGC GCTCCTGCAT CTGCCTCCCC ATATTCCTCG GACACCACAC 120  
 65 CCTGCTGCTT TGCCTACATT GCCCCCCCAC TGCCCCGTGC CCACATCAAG GAGTATTTCT 180  
 ACACCACTGG CAAGTGCTCC AACCCAGCAG TCGTCTTTGT CACCCGAAAG AACCGCCAAG 240  
 TGTGTGCCAA CCCAGAGAAG AAATGGGTTC GGGAGTACAT CAACTCTTTG GAGATGAGCT 300

-52-

AGGATGGAGA GTCCTTGAAC CTGAACTTAC ACAAATTTGC CTGTTTCTGC TTGCTCTTGT 360  
 CCTAGCTTGG GAGGCTTCCC CTCACTATCC TACCCACCCC GCTCCTTGAA GGGCCCAGAT 420  
 5 TCTGACCACG ACGAGCAGCA GTTACAAAAA CCTTCCCCAG GCTGGACGTG GTGGCTCAGC 480  
 CTTGTAATCC CAGCACTTTG GGAGGCCAAG GTGGGTGGAT CACTTGAGGT CAGGAGTTCG 540  
 AGACAGCCTG GCCAACATGA TGAAACCCCA TGTGTACTAA AAATACAAAA AATTAGCCGG 600  
 10 GCGTGGTAGC GGGCGCCTGT AGTCCCAGCT ACTCGGGAGG CTGAGGCAGG AGAATGGCGT 660  
 GAACCCGGGA GCGGAGCTTG CAGTGAGCCG AGATCGCGCC ACTGCACTCC AGCCTGGGCG 720  
 15 ACAGAGCGAG ACTCCGTCTC AAAAAAAAAA AAAAAAAAAA AAAAAATACA AAAATTAGCC 780  
 GCGTGGTGGC CCACGCCTGT AATCCCAGCT ACTCGGGAGG CTAAGGCAGG AAAATTGTTT 840  
 GAACCCAGGA GGTGGAGGCT GCAGTGAGCT GAGATTGTGC CACTTCACTC CAGCCTGGGT 900  
 20 GACAAAGTGA GACTCCGTCA CAACAACAAC AACAAAAAGC TTCCCCAACT AAAGCCTAGA 960  
 AGAGCTTCTG AGGCGCTGCT TTGTCAAAG GAAGTCTCTA GGTCTGAGC TCTGGCTTTG 1020  
 25 CCTTGGCTTT GCAAGGGCTC TGTGACAAGG AAGGAAGTCA GCATGCCTCT AGAGGCAAGG 1080  
 AAGGGAGGAA CACTGCACTC TTAAGCTTCC GCCGTCTCAA CCCCTCACAG GAGCTTACTG 1140  
 GCAAACATGA AAAATCGGGG 1160

## (2) INFORMATION FOR SEQ ID NO:38:

- (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 97 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

45 Met Arg Ile Ser Ala Thr Leu Leu Cys Leu Leu Leu Ile Ala Ala Ala  
 1 5 10 15  
 Phe Ser Ile Gln Val Trp Ala Gln Pro Asp Gly Pro Asn Ala Ser Thr  
 20 25 30  
 50 Cys Cys Tyr Val Lys Lys Gln Lys Ile Pro Lys Arg Asn Leu Lys Ser  
 35 40 45  
 55 Tyr Arg Arg Ile Thr Ser Ser Arg Cys Pro Trp Glu Ala Val Ile Phe  
 50 55 60  
 Lys Thr Lys Lys Gly Met Glu Val Cys Arg Glu Ala His Gln Lys Trp  
 65 70 75 80  
 60 Val Glu Glu Ala Ile Ala Tyr Leu Asp Met Lys Thr Pro Thr Pro Lys  
 85 90 95  
 Pro

65

-53-

## (2) INFORMATION FOR SEQ ID NO:39:

## (1) SEQUENCE CHARACTERISTICS:

- 5 (A) LENGTH: 593 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

10

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

ACTGAAGCCA GCTCTCTCAC TCTCTTCTC CACCATGAGG ATCTCTGCCA CGCTTCTGTG 60  
15 CCTGCTGCTC ATAGCCGCTG CTTTCAGCAT CCAAGTGTGG GCCCAACCAG ATGGGCCCAA 120  
TGCATCCACA TGCTGCTATG TCAAGAAAC AAGATCCCC AAGAGGAATC TCAAGAGCTA 180  
DO CAGAAGGATC ACCAGTAGTC GGTGTCCCTG GGAAGCTGTT ATCTTCAAGA CAAAGAAGGG 240  
20 CATGGAAGTC TCTCGTGAAG CCCATCAGAA GTGGGTCGAG GAGGCTATAG CATACTTAGA 300  
CATGAAAACC CCAACTCCAA AGCCTTGAAG AAATGTGCCT GAACAGAAAC CAACCTAGGA 360  
25 GCCAAGAAGC AAAAATTCCT CACCGCTGTT CTTTCTGAGA ACTGTTGATG AAATGTGTTG 420  
ATCACGGTCC TAAGGGATAG GAGCTGTCTG TAGGAATGTG AAACAGTCAC GCCTAAGGAA 480  
TGGTCTTTAA GTTATTAATA TTTTATTATA ATTAGCCATG TACTTTGGTG TGATTTGAAT 540  
30 GTAAAGCTCT GGAGACCTCA TGCACTTTA ACATTGTGTT AGCTGCAGAA TTC 593

## (2) INFORMATION FOR SEQ ID NO:40:

35

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 72 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

40

## (ii) MOLECULE TYPE: peptide

45

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

Asp Ser Val Ser Ile Phe Ile Thr Cys Cys Phe Asn Val Ile Asn Arg  
1 5 10 15  
50 Lys Ile Pro Ile Gln Arg Leu Glu Ser Tyr Thr Arg Ile Thr Asn Ile  
20 25 30  
Gln Cys Pro Lys Glu Ala Val Ile Phe Lys Thr Gly Lys Glu Val Cys  
35 40 45  
55 Ala Asp Pro Lys Glu Arg Trp Val Arg Asp Ser Met Lys His Lys Asp  
50 55 60  
60 Gln Ile Phe Gln Asn Leu Lys Pro  
65 70

## (2) INFORMATION FOR SEQ ID NO:41:

65

## (1) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

-54-

(ii) MOLECULE TYPE: peptide

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

Asn	Leu	Ala	Lys	Gly	Lys	Glu	Glu	Ser	Leu	Asp	Ser	Asp	Leu	Cys
1				5					10				15	

10 (2) INFORMATION FOR SEQ ID NO:42:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 32 amino acids

(B) TYPE: amino acid

15 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

20

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

Cys	Asn	Gln	Val	Glu	Val	Ile	Ala	Thr	Leu	Lys	Asp	Gly	Arg	Lys	Ile
1				5					10					15	

Cys	Leu	Asp	Pro	Asp	Ala	Pro	Arg	Ile	Lys	Lys	Ile	Val	Gln	Lys	Lys
			20					25					30		

30 (2) INFORMATION FOR SEQ ID NO:43:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 96 base pairs

(B) TYPE: nucleic acid

35 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

40 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

TGCAACCAAG TCGAAGTGAT AGCCACACTG AAGGATGGGA GGAAAATCTG CCTGGACCCA 60

45 GATGCTCCCA GAATCAAGAA AATTGTACAG AAAAAA 96

## CLAIMS

We claim:

1. A method of screening for AHA compounds comprising the steps of:
  - a) contacting a compound with radiolabeled heparin/heparan sulfate and  
5 heparanase;
  - b) maintaining the compounds in contact with the radiolabeled heparin/heparan sulfate and heparanase for a time and under such conditions sufficient to allow inhibition of heparanase activity;
  - c) detecting inhibition of heparanase activity (a compound that gives 50%  
10 inhibition at a concentration of 1  $\mu$ M or less); and
  - d) selecting compounds that inhibit heparanase activity.
2. A method according to claim 1 wherein the heparanase is recombinant.
- 15 3. A heparanase having an isoelectric point of less than 5.5 and possessing activity greater than 20 units heparanase activity per  $\mu$ g protein.
4. A heparanase according to Claim 3, having an isoelectric point of about 4.8 - 5.1.
- 20 5. A heparanase purified to apparent homogeneity, as in claim 3, prepared in the presence of reducing conditions and activated with transglutaminase, having an amino acid sequence selected from the group consisting of SEQ. ID NO: 1, SEQ. ID NO: 3, SEQ. ID NO: 5 or SEQ. ID NO: 7.
- 25 6. A heparanase purified to apparent homogeneity, as in claim 3, prepared in the presence of reducing conditions, having an amino acid sequence of SEQ. ID NO: 1.
7. A heparanase purified to apparent homogeneity, as in claim 3, prepared in the presence of reducing conditions and activated with transglutaminase, having an amino acid sequence of  
30 SEQ. ID NO: 3.
8. A heparanase according to Claim 4, purified to apparent homogeneity, prepared in the presence of reducing conditions and activated with transglutaminase.
- 35 9. A heparanase, as in claim 3, prepared by recombinant means, activated with transglutaminase and having an amino acid sequence selected from the group consisting of SEQ.



ID NO: 1, SEQ. ID NO: 3, SEQ. ID NO: 5 or SEQ. ID NO: 7.

10. A heparanase, as in claim 3, prepared by recombinant means, activated with transglutaminase and having an amino acid sequence of SEQ. ID NO: 1.

5

11. A heparanase, as in claim 3, prepared by recombinant means, activated with transglutaminase and having an amino acid sequence of SEQ. ID NO: 3.

12. A heparanase purified to apparent homogeneity, as in claim 3, prepared in the presence of reducing conditions and activated with transglutaminase, and having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 12, SEQ. ID. NO: 14; SEQ. ID. NO: 16, SEQ. ID. NO: 18, SEQ. ID. NO: 20, SEQ. ID. NO: 22 and SEQ. ID. NO: 24.

13. A heparanase purified to apparent homogeneity, as in claim 3, prepared in the presence of reducing conditions and activated with transglutaminase, and having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 26, SEQ. ID. NO: 28, SEQ. ID. NO: 30, SEQ. ID. NO: 32; SEQ. ID. NO: 34, SEQ. ID. NO: 36, SEQ. ID. NO: 38 and SEQ. ID. NO: 40.

14. A method according to claim 1 wherein the heparanase is purified to apparent homogeneity, prepared in the presence of reducing conditions and activated with transglutaminase, and having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 12, SEQ. ID. NO: 14; SEQ. ID. NO: 16, SEQ. ID. NO: 18, SEQ. ID. NO: 20, SEQ. ID. NO: 22 and SEQ. ID. NO: 24.

15. A method according to claim 1 wherein the heparanase is purified to apparent homogeneity, prepared in the presence of reducing conditions and activated with transglutaminase, and having an amino acid sequence selected from the group consisting of SEQ. ID. NO: 26, SEQ. ID. NO: 28, SEQ. ID. NO: 30, SEQ. ID. NO: 32; SEQ. ID. NO: 34, SEQ. ID. NO: 36, SEQ. ID. NO: 38 and SEQ. ID. NO: 40.

30

16. A method according to claim 1 wherein the heparanase is purified to apparent homogeneity, prepared in the presence of reducing conditions and activated with transglutaminase, and having an amino acid sequence selected from the group consisting of SEQ. ID NO: 1, SEQ. ID NO: 3, SEQ. ID NO: 5 or SEQ. ID NO: 7.

35

17. A peptide having an amino acid sequence of SEQ. ID. NO: 42.

# INTERNATIONAL SEARCH REPORT

International Application No  
PCT/US 94/08207

## A. CLASSIFICATION OF SUBJECT MATTER

IPC 6 C12Q1/34 C12N9/24 C12N9/96 C07K14/47

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 6 C12Q C12N A61K C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	ANALYTICAL BIOCHEMISTRY, vol.157, no.1, 15 August 1986, NEW YORK US pages 162 - 171 MOTOWO NAKAJIMA ET AL. 'A Solid-Phase Substrate of Heparanase: Its Application to Assay of Human Melanoma for Heparan Sulfate Degradative Activity' see the whole document ---	1,2
A	JOURNAL OF BIOLOGICAL CHEMISTRY, vol.259, no.4, 25 February 1984, BALTIMORE, MD US pages 2283 - 2290 MOTOWO NAKAJIMA ET AL. 'Metastatic Melanoma Cell Heparanase' see the whole document	1,2
A	see page 2289, left column, line 63 - right column, line 31 ---	3,4
-/--		

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

### \* Special categories of cited documents:

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*B\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*A\* document member of the same patent family

Date of the actual completion of the international search

9 December 1994

Date of mailing of the international search report

21.12.94

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax (+31-70) 340-3016

Authorized officer

Döpfer, K-P

# INTERNATIONAL SEARCH REPORT

Int'l Application No  
PCT/US 94/08207

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO,A,91 02977 (HADASSAH MEDICAL ORGANIZATION) 7 March 1991 cited in the application see page 23, paragraph 6.1.5 - page 26, paragraph 6.1.7; claims 1-3 ---	3-16
A	FASEB JOURNAL, vol.5, no.15, December 1991, BETHESDA, MD US pages 3071 - 3076 CHARLES S. GREENBERG ET AL. 'Transglutaminases: multifunctional cross-linking enzymes that stabilize tissues' cited in the application see the whole document ---	3-16
A	JOURNAL OF BIOLOGICAL CHEMISTRY, vol.265, no.28, 5 October 1990, BALTIMORE, MD US pages 17180 - 17188 ELEONORA CORDELLA-MIELE ET AL. 'A Novel Transglutaminase-mediated Post-translational Modification of Phospholipase A2 Dramatically Increases Its Catalytic Activity' cited in the application see the whole document ---	3-16
A	WO,A,93 09794 (UNIVERSITY OF PENNSYLVANIA) 27 May 1993 see especially Sequence Description: SEQ ID NO: 1 see page 17, line 1 - line 16 -----	17

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/US 94/08207

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9102977	07-03-91	AU-B- 654804	24-11-94
		AU-A- 6336490	03-04-91
		CA-A- 2065744	24-02-91
		EP-A- 0487627	03-06-92
		JP-T- 5504047	01-07-93
		US-A- 5362641	08-11-94
WO-A-9309794	27-05-93	AU-A- 3071492	15-06-93